

Méthodes de Monte Carlo pour la finance

DEA Statistiques et Modèles Aléatoires en Économie et
Finance

Université Paris VII

Emmanuel TEMAM
Université Paris VII, LPMA
temam@math.jussieu.fr

Avril 2004

Table des matières

1	Introduction	5
1.1	Présentation	5
1.2	Rappels sur les méthodes de Monte Carlo	6
1.3	Notations	7
2	Schémas d'approximation d'EDS	9
2.1	Simulation du mouvement Brownien et des processus rattachés	9
2.1.1	Simulation du modèle de Black Scholes	9
2.1.2	Une construction du mouvement Brownien	10
2.2	Discrétisation d'EDS	11
2.2.1	Le schéma d'Euler	11
2.2.2	Le schéma de Milshtein	13
2.2.3	Schémas d'ordre supérieur	16
2.3	Méthodes spéciales pour les options exotiques	17
2.3.1	Options asiatiques	17
2.3.2	Options sur maximum	21
2.3.3	Options barrières	24
3	Réduction de variance	29
3.1	Fonctions d'importance	29
3.1.1	Un exemple en finance	29
3.1.2	Généralités	30
3.1.3	Théorème de Girsanov et fonctions d'importance pour les diffusions	31
3.2	Variables antithétiques	32
3.3	Variables de contrôles	34
3.3.1	Théorie	34
3.3.2	Exemples	35
3.3.3	Valeur moyenne et conditionnement	38
4	Calcul des sensibilités	41
4.1	Rappels sur les sensibilités	41
4.2	Approche par différences finies	42
4.3	Amélioration des techniques pour Monte Carlo	42
4.3.1	Notion de processus tangent	42
4.3.2	Applications aux calculs de couverture	44
4.4	Calcul de Malliavin	46

4.4.1	Introduction au calcul de Malliavin	47
4.4.2	Applications aux calculs des sensibilités	51
5	Les options américaines	55
5.1	Introduction	55
5.2	L'algorithme de Longstaff Schwartz	57
5.2.1	Une équation de programmation dynamique pour τ^*	57
5.2.2	La méthode de régression	57
5.2.3	Du Monte Carlo pour le problème de régression	58
5.3	Approximation d'espérances conditionnelles	59
5.3.1	Calcul de Malliavin : compléments	60
5.4	Application à l'évaluation d'options américaines	67

Chapitre 1

Introduction

1.1 Présentation

Dans un modèle stochastique, les prix des actifs sont représentés comme la solution d'une équation différentielle stochastique. Le problème essentiel posé les mathématiques financières est d'évaluer les produits dérivés sur ces actifs.

Dans la plupart des cas, le payoff de ces produits est donné par une fonction de l'actif sous-jacent à une (ou des) date(s) future(s). Le prix, dans un modèle complet, est alors l'espérance sous l'unique probabilité risque neutre du payoff actualisé. L'intérêt numérique est donc de calculer cette espérance de la manière la plus efficace et la plus rapide possible.

Il existe de nombreux types de méthodes pour cela :

- Ces espérances peuvent à l'aide de la formule de Feynman-Kac s'écrire comme la solutions d'équations aux dérivées partielles. Ainsi, on peut utiliser les méthodes existantes de type éléments finis ou différences finies.
- Il existe également les méthodes dites d'arbre, qui consistent à approcher la solution de l'équation différentielle stochastique par une chaîne de Markov discrète.
- Enfin, les méthodes de Monte Carlo sur lesquelles nous nous concentrerons. Ces méthodes nécessitent de savoir simuler l'EDS du sous-jacent. Souvent, il faut recourir à des schémas numériques, ce sera l'objet du chapitre 1. De plus, même cette discrétisation effectuée, il peut s'avérer que la méthode ne soit pas efficace, par exemple lorsque la variance est trop élevée. Les méthodes de réduction de variance permettent d'éviter ce genre de difficulté (voir chapitre 2).

La finance pose d'autres problèmes nettement plus ardues et qui sont encore aujourd'hui l'objet de développement constant :

- l'évaluation du prix n'est pas le plus important et même un prix ne sert à rien si on ne donne pas une stratégie. L'évaluation des grecques (les dérivées du prix par rapport au point initial ou à la volatilité) est un problème complexe du au fait qu'il s'agit d'une différentielle (voir chapitre 3).
- En fait, sur le marché, on ne dispose pas de volatilité. Il faut évaluer les paramètres du modèle à l'aide de prix de certaines options tellement liquides qu'elles sont cotées directement. Ce type de problème s'appelle la calibration.
- Enfin, il existe des options qui peuvent être exercées à n'importe quel instant. Ce sont les options américaines et leurs évaluations posent des problèmes plus complexes (voir chapitre

4).

1.2 Rappels sur les méthodes de Monte Carlo

Les méthodes de Monte Carlo sont nées grâce à Louis Buffon mathématicien du $XVII^{eme}$ siècle. Il a prouvé que si l'on se donne un parquet dont les lattes sont écartées d'une longueur d , et que l'on jette toujours de la même hauteur une aiguille de longueur l la probabilité que l'aiguille touche deux stries est de ????. Il a ainsi entrevu une possibilité pour calculer le nombre π numériquement. En effet, en lançant un grand nombre d'aiguilles et en comptant celles qui touchaient deux stries il détenait une valeur approximative de la probabilité. Cette expérience repose sur un théorème : **la loi forte des grands nombres**.

Théorème 1.1 (Loi forte des grands nombres). *Soit $(X_i, i \geq 1)$ une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) telles que $\mathbb{E}(|X_1|) < +\infty$, alors*

$$\lim_{n \rightarrow +\infty} \left(\frac{1}{n} \sum_{i=1}^n X_i \right) = \mathbb{E}(X_1).$$

Ce théorème permet d'utiliser des algorithmes probabilistes pour calculer n'importe quelle espérance. Considérons le cas où les variables aléatoires suivent une loi uniforme sur $[0, 1]$. Alors un générateur aléatoire permet de simuler cette loi. Pour approcher son espérance, il suffit donc de simuler n fois cette loi et de faire la moyenne des simulations. C'est aussi pour cela que le terme de gauche est souvent appelé moyenne arithmétique et on le notera \bar{X}_n .

Il faut remarquer également que cette méthode permet de calculer un certain nombre d'intégrales. Soit $I = \int_0^1 f(u)du$ où f est une fonction intégrable. On voit immédiatement que I se réécrit comme $I = \mathbb{E}[(f(U))]$ où U est une variable aléatoire de loi $\mathcal{U}_{[0,1]}$. En appliquant l'algorithme précédant à $X_i = f(U_i)$ on calcule I .

Le problème se pose maintenant d'évaluer théoriquement l'efficacité de cette méthode. Il existe des théorèmes qui permettent de calculer la vitesse de convergence d'une méthode de Monte Carlo.

Théorème 1.2 (Théorème de la Limite Centrale). *Soit $(X_i, i \geq 1)$ une suite de variables aléatoires, indépendantes et identiquement distribuées telles que $\mathbb{E}(X_i^2) < +\infty$. On pose $\sigma^2 = \text{Var}(X_1^2)$, alors*

$$\frac{\sqrt{n}}{\sigma^2} \left(\mathbb{E}[(X_1)] - \frac{1}{n}(X_1 + \dots + X_n) \right) \underset{\mathcal{L}}{\rightsquigarrow} G \quad \text{où } G \rightsquigarrow \mathcal{N}(0, 1).$$

Remarque 1.1. – *Contrairement aux méthodes non probabilistes, la méthode de Monte Carlo, utilisée pour calculer des intégrales, ne nécessite pas de fonctions régulières.*

- *L'erreur que donne le Théorème de la Limite Centrale n'est pas bornée car le support de la gaussienne est \mathbb{R} .*
- *De ce théorème on peut déduire des intervalles de confiance qui sont des indicateurs importants de la performance d'une méthode de Monte Carlo.*

Il faut également noter que l'évaluation de la variance est cruciale pour mesurer l'efficacité de la méthode. En effet, en supposant que celle-ci n'est pas aléatoire, on peut réécrire le

théorème de la limite centrale comme :

$$\sqrt{n} \left(\mathbb{E}[(X_1)] - \frac{1}{n}(X_1 + \dots + X_n) \right) \xrightarrow{\mathcal{L}} G \quad \text{où } G \rightsquigarrow \mathcal{N}(0, \sigma^2).$$

Ainsi, si on veut construire des intervalles de confiance, leurs largeurs dépend du nombre de simulations et de la variance. Pour résumer, l'efficacité des méthodes de Monte Carlo s'écrit

$$\frac{\sigma^2}{\sqrt{n}}.$$

Pour évaluer la variance, on dispose d'estimateurs. Soit (X_1, \dots, X_n) , une suite de variables aléatoires vérifiant les hypothèses du théorème de la limite centrale. L'estimateur standard de la variance est :

$$\widetilde{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

Cet estimateur est construit par la loi forte des grands nombres. En effet, $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. En approchant, les deux espérances par leurs moyennes arithmétiques on retombe sur $\widetilde{\sigma}_n$. Cependant, on peut montrer que cette estimateur est biaisé $\mathbb{E}[\widetilde{\sigma}_n^2] = \frac{n}{n-1} \text{Var}(X)$. L'estimateur sans biais est donc :

$$\overline{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

Pour plus de détails sur les méthodes de Monte Carlo on pourra consulter [40, 44]. Par ailleurs, nous aurons besoin d'un théorème qui lie les méthodes probabilistes et analytiques :

Théorème 1.3. (i) *Supposons que b et a sont lipschitziennes et qu'il existe une solution $u \in C_p^{1,2}$ à l'équation parabolique*

$$\begin{aligned} 0 &= \frac{\partial u}{\partial t} + \sum_{j=1}^d b^j \frac{\partial u}{\partial x^j} + \frac{1}{2} \sum_{i,j=1}^d (aa^*)^{ij} \frac{\partial^2 u}{\partial x^i \partial x^j}, \quad \text{sur } [0, T] \times \mathbb{R}^d \\ g &= u(T, \cdot) \quad \text{sur } \mathbb{R}^d, \end{aligned} \tag{1.1}$$

alors

$$u(t, x) = \mathbb{E}[g(X_T) | X_t = x] \quad \text{et} \quad g(X_T) = \mathbb{E}[g(X_T)] + \int_0^T \nabla_x u(t, X_t) a(X_t) dW_t. \tag{1.2}$$

(ii) *Si $a \in C_b^4$, $b \in C_b^2$ et $g \in C_p^4$, alors la réciproque est vraie et $u \in C_p^4$.*

1.3 Notations

Dans la suite on notera S_t la solution de l'équation différentielle stochastique, X_t la solution d'une équation différentielle stochastique générale.

On commence par introduire quelques notations. On identifiera un élément x de \mathbb{R}^d au vecteur colonne associé de coordonnées x^j , $j = 1, \dots, d$. On notera I_d la matrice identité de \mathbb{M}^d et e_d le vecteur unité de \mathbb{R}^d . La norme euclidienne sur \mathbb{R}^d ou \mathbb{M}^d sera notée $\|\cdot\|$. Pour un élément a de \mathbb{M}^d , on notera a^i le vecteur ligne correspondant à sa i -ème ligne et a^j le vecteur colonne correspondant à sa j -ème colonne. La transposée de $a \in \mathbb{M}^d$ sera notée a^* . Pour un ensemble de point $(x^i)_{i=1}^d$, on notera $\text{Vect}[(x^i)_{i=1}^d]$ le vecteur colonne de composantes x^i . Pour $x \in \mathbb{R}^d$, on notera $\text{diag}[x]$ la matrice diagonale de \mathbb{M}^d dont les éléments diagonaux sont les x^i . Pour une fonction $f : \mathbb{R}^d \mapsto \mathbb{R}^n$, on notera ∇f sa matrice Jacobienne, i.e. $(\nabla f)^{ij} = \partial f^i / \partial x^j$. Si $f : (\mathbb{R}^d)^k \mapsto \mathbb{R}^n$, on notera $\nabla_\ell f$, la matrice Jacobienne obtenue en dérivant par rapport à sa ℓ -ème composante vectorielle, i.e. $(\nabla_\ell f(x_1, \dots, x_k))^{ij} = \partial f^i(x_1, \dots, x_k) / \partial x_\ell^j$. On écrira parfois simplement $\nabla_{x_\ell} f$. On notera C_p^k (resp. C_b^k) l'ensemble des fonctions C^k à croissance polynômiale (resp. bornées) dont les dérivées sont également à croissance polynômiale (resp. bornées). Lorsque la fonction est seulement définie sur A , on notera $C^k(A)$, $C_p^k(A)$ et $C_b^k(A)$. $\mathcal{N}(m, \Sigma)$ désignera la loi normale de moyenne m et de matrice de variance-covariance Σ . On notera souvent par $C > 0$ une constante générique qui peut changer de valeur d'une ligne à l'autre.

Chapitre 2

Schémas d'approximation d'EDS

Le but de cette partie est de décrire les méthodes utilisées pour la simulation trajectorielle d'un processus donné. Cette simulation est nécessaire lorsque l'on veut calculer une option dépendant de la trajectoire (options asiatiques, barrières, etc...) dont le prix peut s'écrire :

$$\mathbb{E}[f(X_s, s \leq T)],$$

où φ est la fonctionnelle du processus X .

2.1 Simulation du mouvement Brownien et des processus rattachés

Nous commençons par la simulation du mouvement Brownien. La simulation approximative d'une trajectoire du mouvement Brownien est un algorithme simple. En effet, soit h un intervalle de discrétisation. Soit $(G_k, 0 \leq k \leq N)$ une réalisation d'une famille de variables aléatoires normales de moyenne 0 et de variance 1. Pour obtenir une réalisation de $(W_{ph}, 1 \leq p \leq N)$, il suffit de calculer :

$$W_{ph}^h = \sqrt{h} \sum_{1 \leq k \leq p} G_k.$$

On peut remarquer que dans ce cas la loi de l'approximation $(W_{ph}^h, 1 \leq p \leq N)$ est identique à la loi du vecteur $(W_{ph}, 1 \leq p \leq N)$: **il n'y a donc pas d'erreur de discrétisation.**

2.1.1 Simulation du modèle de Black Scholes

Soit

$$S_t = x \exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma W_t\right).$$

En utilisant le paragraphe précédent, on construit une simulation exacte en loi du vecteur $(S_{ph}, 1 \leq p \leq N)$ en posant :

$$S_{ph}^h = x \exp\left(\left(r - \frac{\sigma^2}{2}\right)ph + \sigma W_{ph}^h\right).$$

Il faut remarquer que l'on peut obtenir une simulation exacte quand la volatilité σ est une fonction du temps puisque

$$\int_{ph}^{(p+1)h} \sigma_s dW_s,$$

reste une suite de variable aléatoire gaussienne de moyenne 0 et de variance $\int_{ph}^{(p+1)h} \sigma_s^2 ds$.

2.1.2 Une construction du mouvement Brownien

Dans certains cas nous avons besoin d'une méthode de simulation qui peut être raffinée en certains points de la trajectoire. Cela est une des caractéristiques de la méthode suivante. Il faut noter que cet algorithme donne aussi une construction d'une trajectoire du mouvement Brownien.

Cette simulation repose sur le calcul de la loi conditionnelle de $W_{\frac{t+s}{2}}$ sachant W_s et W_t . Si $s \leq t$, le vecteur $(W_s, W_{\frac{t+s}{2}}, W_t)$ est un vecteur gaussien. On peut donc trouver deux nombres α et β tels que la variable aléatoire $Z_{\alpha,\beta}$ définie par

$$Z_{\alpha,\beta} = W_{\frac{t+s}{2}} - \alpha W_s - \beta W_t,$$

soit indépendante de (W_s, W_t) . Le vecteur $(Z_{\alpha,\beta}, W_s, W_t)$ est encore un vecteur gaussien de moyenne nulle. Donc, $Z_{\alpha,\beta}$ est indépendante de W_s et de W_t si et seulement si

$$\text{Cov}(Z_{\alpha,\beta}, W_s) = 0 \quad \text{et} \quad \text{Cov}(Z_{\alpha,\beta}, W_t) = 0.$$

Comme

$$\text{Cov}(Z_{\alpha,\beta}, W_s) = \mathbb{E} \left(W_{\frac{s+t}{2}} W_s \right) - \alpha \mathbb{E}(W_s^2) - \beta \mathbb{E}(W_t^2),$$

et que $\forall t$ et $\forall s$

$$\mathbb{E}(W_s W_t) = t \wedge s = \inf(s, t),$$

on obtient

$$\text{Cov}(Z_{\alpha,\beta}, W_s) = s - \alpha s - \beta s,$$

et :

$$\text{Cov}(Z_{\alpha,\beta}, W_t) = \frac{s+t}{2} - \alpha s - \beta t.$$

Il est clair que les deux covariances s'annulent si et seulement si $\alpha = \beta = 1/2$. De plus, comme $Z_{\alpha,\beta}$ est une combinaison linéaire d'éléments d'un vecteur gaussien, c'est une variable gaussienne. Pour déterminer entièrement sa loi, il suffit de calculer sa moyenne et sa variance. Si $Z = Z_{1/2,1/2}$ alors

$$Z = W_{\frac{t+s}{2}} - \frac{1}{2}(W_s + W_t).$$

Ainsi $\mathbb{E}(Z) = 0$ et

$$\begin{aligned} \text{Var}(Z) &= \mathbb{E}(Z^2) = \mathbb{E} \left\{ \left(W_{\frac{s+t}{2}} - \frac{s+t}{2}(W_s + W_t) \right)^2 \right\} \\ &= \frac{t+s}{2} + \frac{t}{4} + \frac{s}{4} - \frac{t+s}{2} - s + \frac{s}{2} = \frac{1}{4}(t-s). \end{aligned} \quad (2.1)$$

Z peut donc s'écrire :

$$Z = \frac{1}{2}\sqrt{t-s}G_{t,s},$$

où $G_{t,s}$ est une variable aléatoire gaussienne de moyenne nulle et de variance 1, indépendante de (W_s, W_t) .

Remarque 2.1. On peut montrer que $G_{t,s}$ est indépendante de $(W_u, u \leq s)$ et $(W_u, u \geq t)$.

On peut résumer ce raisonnement comme suit :

$$\begin{cases} W_{\frac{t+s}{2}} = \frac{1}{2}(W_t + W_s) + \frac{1}{2}\sqrt{t-s}G_{t,s} \\ G_{t,s} \text{ est indépendant de } (W_u, u \leq s) \text{ et } (W_u, u \geq t), \end{cases} \quad (2.2)$$

Cela permet d'envisager une procédure de simulation (ou de construction) de la trajectoire de $(X_s, 0 \leq s \leq 1)$

1. Calculer X_1 avec une variable $\mathcal{N}(0, 1)$,
2. Utiliser 2.2, avec $s = 0$ et $t = 1$ pour simuler $X_{1/2}$,
3. Utiliser 2.2, avec $s = 0$ et $t = 1/2$ pour simuler $X_{1/4}$,
4. Utiliser 2.2, avec $s = 1/2$ et $t = 1$ pour simuler $X_{3/4}$,
5. et ainsi de suite!

Cette méthode peut être utile pour calculer les chemins du processus de Wiener sur des "meshs" dont la taille se réduit de plus en plus.

2.2 Discrétisation d'EDS

La méthode de Monte Carlo utilisée pour approcher $\mathbb{E}[f(X_T)]$ suppose que l'on sait simuler la loi de la variable aléatoire X_T . Or, en général, on ne peut pas résoudre explicitement l'équation différentielle stochastique associée au processus X . Par exemple, la plupart des modèles de taux d'intérêt conduisent à des équations qui n'ont pas de solutions explicites. De plus, même si on trouve une solution, celle-ci peut être trop complexe pour être simulée directement. Ainsi, par analogie avec les EDP, il est naturel de chercher à simuler une solution à partir de l'équation elle-même en utilisant des schémas d'approximation. Ainsi, la méthode de Monte Carlo va consister en l'approximation de $\mathbb{E}[f(X_T)]$ par $\sum_i f(\bar{X}_t^n)$ où \bar{X}^n est le schéma. L'erreur de ces méthodes a deux causes : une erreur statistique et une erreur liée à la discrétisation.

2.2.1 Le schéma d'Euler

Notre but est de trouver un schéma approchant la solution d'une équation différentielle stochastique. Soit $(X_t, t \geq 0)$ le processus d -dimensionnel solution de

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \quad (2.3)$$

où $(W_t, t \geq 0)$ est un mouvement Brownien r -dimensionnel.

Le schéma le plus simple est celui d'Euler. Soit n le nombre d'intervalles de discrétisation choisi et $h = T/n$. La solution exacte vérifie :

$$X_h = X_0 + \int_0^h b(X_s)ds + \int_0^h \sigma(X_s)dW_s.$$

En se fondant sur la définition d'un processus d'Itô, une approximation naturelle est X_h définie par :

$$X_h \simeq X_0 + b(X_0)h + \sigma(X_0)(W_h - W_0).$$

En procédant par récurrence, on obtient le schéma d'Euler pour l'équation différentielle stochastique (2.3) :

$$\begin{aligned} \bar{X}_0^n &= X_0, \\ \bar{X}_{t_{k+1}}^n &= \bar{X}_{t_k}^n + b(\bar{X}_{t_k}^n)h + \sigma(\bar{X}_{t_k}^n)(W_{t_{k+1}} - W_{t_k}). \end{aligned} \quad (2.4)$$

Comme l'on est en dimension supérieur à 1, il faut faire attention aux produits matriciels :

$$\begin{aligned} \bar{X}_0^n &= X_0, \\ \bar{X}_{t_{k+1}}^{n,i} &= \bar{X}_{t_k}^{n,i} + b^i(\bar{X}_{t_k}^n)h + \sum_{j=1}^r \sigma^{i,j}(\bar{X}_{t_k}^n)(W_{t_{k+1}}^j - W_{t_k}^j) \quad 1 \leq i \leq d \end{aligned} \quad (2.5)$$

Ce schéma est une généralisation naturelle aux EDS des schémas d'Euler utilisés pour les équations différentielles ordinaires. La simulation d'un schéma d'Euler est extrêmement simple puisqu'il suffit de simuler la variable gaussienne $W_h - W_0 = W_h$.

Le théorème suivant donne les résultats de convergence connus pour le schéma d'Euler. Nous renvoyons au polycopié sur les méthodes de Monte Carlo pour les démonstrations (voir aussi [22, 23, 55]).

Théorème 2.1. *Soient b et σ deux fonctions lipschitziennes. Soit $(W_t, t \geq 0)$ un mouvement Brownien r -dimensionnel. On note $(X_t, t \geq 0)$ l'unique solution de*

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, X_0 = x,$$

et par $(\bar{X}_{kh}^n, k \geq 0)$ la suite de variables aléatoires définies par l'équation (2.4). Alors, pour tous $q \geq 1$:

– **Convergence forte :**

$$\mathbb{E} \left(\sup_{k, kh \leq T} |\bar{X}_{kh}^n - X_{kh}|^{2q} \right) \leq Ch^q.$$

De plus, pour tous $\alpha < 1/2$, presque sûrement

$$\lim_{h \rightarrow 0} h^\alpha \sup_{k, kh \leq T} |\bar{X}_{kh}^n - X_{kh}| = 0,$$

- **Convergence faible** : Si b et σ sont des fonctions C^4 avec des dérivées bornées jusqu'à l'ordre 4 et f est une fonction C^4 à croissance polynômiale avec des dérivées bornées jusqu'à l'ordre 4 si $h = T/N$, alors il existe une constante C_T telle que

$$|\mathbb{E}(f(X_T)) - \mathbb{E}(f(\bar{X}_T^n))| \leq \frac{C_T(f)}{N}.$$

Remarque 2.2. Ce théorème prouve que la vitesse de convergence dans L^2 est de l'ordre de $h^{1/2}$ et que la vitesse de convergence presque sûre est d'ordre $h^{1/2-\epsilon}$, pour tout $\epsilon > 0$. Le deuxième résultat montre que pour des fonctions régulières, la vitesse de convergence en loi du schéma est d'ordre h .

2.2.2 Le schéma de Milshtein

Pour les équations différentielles ordinaires, le schéma d'Euler peut être amélioré par les méthodes de Runge Kutta. Plusieurs schémas d'ordre supérieur ont été proposés pour les EDS. Nous verrons cependant que leur mise en oeuvre reste délicate.

Le plus simple schéma d'ordre 2 est le schéma de Milshtein. Il permet de faire converger à une vitesse supérieure dans les espaces L^p mais est difficile à simuler quand la dimension est strictement plus grande que 1 et converge en loi à la même vitesse que le schéma d'Euler.

Cas de la dimension 1

Nous commençons par construire le schéma de Milshtein quand $d = r = 1$. Ce schéma est défini par $\tilde{X}_0^n = x$ et pour $k \geq 1$:

$$\begin{aligned} \tilde{X}^n((k+1)h) &= \tilde{X}_{kh}^n + b\left(\tilde{X}_{kh}^n\right)h + \sigma\left(\tilde{X}_{kh}^n\right)(W_{(k+1)h} - W_{kh}) \\ &\quad + \sigma'\left(\tilde{X}_{kh}^n\right)\sigma\left(\tilde{X}_{kh}^n\right)\int_{kh}^{(k+1)h}(W_s - W_{kh})dW_s. \end{aligned} \quad (2.6)$$

Remarque 2.3. Il est facile de comprendre comment le nouveau terme apparaît en regardant l'équation

$$dX_t = \sigma(X_t)dW_t.$$

On peut étendre le schéma d'Euler à tous t dans $[t_k, t_{k+1}]$ par interpolation linéaire. Si $t_k = kh$

$$\hat{X}_t^n = \hat{X}_{t_k}^n + \sigma\left(\hat{X}_{t_k}^n\right)(W_t - W_{t_k}).$$

\hat{X}_t^n donne une approximation de X_t sur l'intervalle $[t_k, t_{k+1}]$ qui est meilleure que $\bar{X}^n(t_k)$. On peut espérer que $\sigma(\hat{X}_t^n)$ est une meilleure approximation de $\sigma(X_t)$ que $\sigma(\bar{X}_{t_k}^n)$.

Un bon candidat pour un schéma d'ordre supérieur est

$$\hat{X}_t^n = \hat{X}_{t_k}^n + \int_{t_k}^t \sigma\left(\hat{X}_s^n\right)dW_s.$$

Ce schéma peut être approché en utilisant des développements de Taylor

$$\begin{aligned} \sigma\left(\hat{X}_t^n\right) &= \sigma\left(\hat{X}_{t_k}^n + \sigma\left(\hat{X}_{t_k}^n\right)(W_t - W_{t_k})\right) \\ &\approx \sigma\left(\hat{X}_{t_k}^n\right) + \sigma'\left(\hat{X}_{t_k}^n\right)\sigma\left(\hat{X}_{t_k}^n\right)(W_t - W_{t_k}). \end{aligned}$$

Cela conduit au schéma suivant

$$\widehat{X}_t^n = \widehat{X}_{t_k}^n + \sigma \left(\widehat{X}_{t_k}^n \right) (W_t - W_{t_k}) + \sigma \left(\widehat{X}_{t_k}^n \right) \sigma' \left(\widehat{X}_{t_k}^n \right) \int_{t_k}^t (W_t - W_{t_k}) dW_s.$$

Nous avons retrouvé le schéma de Milshstein quand $b = 0$. Les calculs précédents s'étendent au cas où $b \neq 0$.

D'un point de vue pratique il est important de noter que par la formule d'Itô :

$$\int_{kh}^{(k+1)h} (W_s - W_{kh}) dW_s = \frac{1}{2} \left((W_{(k+1)h} - W_{kh})^2 - h \right).$$

Le schéma de Milshstein se réécrit donc :

$$\begin{aligned} \widetilde{X}_{(k+1)h}^n &= \widetilde{X}_{kh}^n + \left(b \left(\widetilde{X}_{kh}^n \right) - \frac{1}{2} \sigma' \left(\widetilde{X}_{kh}^n \right) \sigma \left(\widetilde{X}_{kh}^n \right) \right) h + \sigma \left(\widetilde{X}_{kh}^n \right) (W_{(k+1)h} - W_{kh}) \\ &\quad + \frac{1}{2} \sigma' \left(\widetilde{X}_{kh}^n \right) \sigma \left(\widetilde{X}_{kh}^n \right) (W_{(k+1)h} - W_{kh})^2. \end{aligned}$$

Ce schéma est facile à mettre en oeuvre car les variables aléatoires $(W_{(k+1)h} - W_{kh}, k \geq 0)$ sont impliquées.

Exemple 2.1. Considérons le cas où $(S_t, t \geq 0)$ est une diffusion log-normale définie par :

$$dS_t = S_t (r dt + \sigma dW_t), S_0 = x.$$

Soit $\Delta W_k = W_{(k+1)h} - W_{kh}$. Le schéma d'Euler s'écrit

$$\widetilde{X}_{(k+1)h}^n = \widetilde{X}_{kh}^n (1 + rh + \sigma \Delta W_k),$$

et le schéma de Milshstein

$$\widetilde{X}_{(k+1)h}^n = \widetilde{X}_{kh}^n \left(1 + \left(r - \frac{1}{2} \sigma^2 \right) h + \sigma \Delta W_k + \frac{1}{2} \sigma^2 (\Delta W_k)^2 \right).$$

Le schéma de Milshstein en dimension supérieure Quand le nombre r de mouvements Brownien impliqués dans l'équation différentielle stochastique est plus grand que 1, le schéma de Milshstein s'écrit comme

$$\begin{aligned} \widetilde{X}_{(k+1)h}^n &= \widetilde{X}_{kh}^n + b \left(\widetilde{X}_{kh}^n \right) h + \sigma \left(\widetilde{X}_{kh}^n \right) (W_{(k+1)h} - W_{kh}) \\ &\quad + \sum_{j,l=1}^p (\partial \sigma_j \sigma_l) \left(\widetilde{X}_{kh}^n \right) \int_{kh}^{(k+1)h} \left(W_s^j - W_{kh}^j \right) dW_s^l, \end{aligned} \tag{2.7}$$

avec, pour $1 \leq i \leq n$

$$(\partial \sigma_j \sigma_l)_i(x) = \sum_{r=1}^n \frac{\partial \sigma_{ij}(x)}{\partial x_r} \sigma_{rl}(x).$$

Ces formules s'obtiennent de la formule de Taylor comme à l'ordre 1.

Remarque 2.4. Le schéma de Milshtein est difficile à implémenter car l'on doit simuler le vecteur

$$\left(W_{(k+1)h}^j - W_{kh}^j, \int_{kh}^{(k+1)h} (W_s^j - W_{kh}^j) dW_s^l \right),$$

pour $1 \leq j \leq p, 1 \leq l \leq p$. Il est facile de voir que la simulation de

$$\int_{kh}^{(k+1)h} (W_s^j - W_{kh}^j) dW_s^l,$$

ne peut faire intervenir que $W_{(k+1)h} - W_{kh}$. Il n'existe pas à ce jour de méthode efficace pour la simulation de ces lois.

Quand $p = 2$, il est équivalent de simuler

$$\left(W_h^1, W_h^2, \int_0^h W_s^1 dW_s^2 - W_s^2 dW_s^1 \right),$$

et la méthode décrite dans [26] est compliquée.

Le schéma de Milshtein peut être utilisé facilement quand il n'est pas nécessaire de simuler $\int_0^h W_s^1 dW_s^2 - W_s^2 dW_s^1$. C'est le cas en dimension 1 ou quand la condition de commutativité suivante est satisfaite

$$(C) \quad \text{Pour tous } j, k \text{ dans } \{1, \dots, p\} \text{ et pour tous } x \in \mathbb{R}^n : \\ \partial \sigma_j(x) \sigma_k(x) = \partial \sigma_k(x) \sigma_j(x).$$

Sous l'hypothèse (C), on peut réécrire le schéma de Milshtein comme

$$\begin{aligned} \tilde{X}_{(k+1)h}^n &= \tilde{X}_{kh}^n + \left(b(\tilde{X}_{kh}^n) - \frac{1}{2} \sum_{j=1}^p (\partial \sigma_j \sigma_j)(\tilde{X}_{kh}^n) \right) h + \sigma(\tilde{X}_{kh}^n) (W_{(k+1)h} - W_{kh}) \\ &\quad + \frac{1}{2} \sum_{j,l=1}^p (\partial \sigma_j \sigma_l)(\tilde{X}_{kh}^n) \times (W_{(k+1)h}^j - W_{kh}^j) (W_{(k+1)h}^l - W_{kh}^l). \end{aligned}$$

On a alors seulement besoin de

$$\left((W_{(k+1)h}^j - W_{kh}^j), k \geq 0, 1 \leq j \leq p \right).$$

Le théorème suivant donne la vitesse de convergence du schéma de Milshtein

Théorème 2.2. On suppose que b et σ sont deux fois continuellement différentiables avec des dérivées bornées. On note $(X_t, t \geq 0)$ l'unique solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, X_0 = x,$$

et par $(\tilde{X}_{kh}^n, k \geq 0)$ la suite de variables aléatoires définies par (2.7). Alors

– **convergence forte :**

$$\text{pour tous } q \geq 1, \sup_{k, kh \leq T} \mathbb{E} \left(\left| X_{kh} - \tilde{X}_{kh}^n \right|^q \right) \leq Ch^q,$$

De plus, pour tous $\alpha < 1$, on a

$$\lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sup_{k, kh \leq T} \left| X_{kh} - \tilde{X}_{kh}^n \right| = 0,$$

- **Convergence en loi :** Si b et σ sont deux fonctions C^4 avec des dérivées bornées jusqu'à l'ordre 4 et si f est une fonction à croissance polynômiale C^4 avec des dérivées bornées jusqu'à l'ordre 4, alors il existe une constante $C_T(f) > 0$ telle que

$$\left| \mathbb{E}(f(X_T)) - \mathbb{E}\left(f(\tilde{X}_T^n)\right) \right| \leq \frac{C_T(f)}{N}.$$

Remarque 2.5. Il faut remarquer que ce résultat ne dépend pas de l'hypothèse de commutativité (C).

Le schéma de Milshstein améliore les vitesses de convergence trajectorielles : il est d'ordre h alors que le schéma d'Euler est d'ordre \sqrt{h} . Néanmoins, la vitesse de convergence pour des fonctions régulières est la même. Or c'est ce type de convergence qui est important en finance pour les méthodes de Monte Carlo. De plus, le schéma de Milshstein nécessite la simulation de termes supplémentaires ce qui nuit à sa vitesse d'exécution. En conséquence, dans la plupart des cas on utilisera le schéma d'Euler.

2.2.3 Schémas d'ordre supérieur

Nous avons vu qu'il était facile d'obtenir des vitesses de convergence trajectorielle d'ordre \sqrt{h} (schéma d'Euler) ou d'ordre h (schéma de Milshstein). Il est naturel d'essayer de construire des schémas d'ordre supérieurs. Toutefois, un résultat de Clark and Cameron (see [16] ou [23]) prouve que, vis à vis de la norme L^2 le schéma d'Euler est optimal dans la classe des schémas n'utilisant que les variables $W_{ph}, p \geq 1$. Si on accepte d'utiliser plus que ces incréments on peut construire des schémas d'ordre arbitraire (voir [39]). Ces schémas sont très peu utilisés en finance.

Pour tout développement supplémentaire : [56, 39, 23, 22], [36], [55] et [39].

Méthodes de Romberg pour Euler et Milshstein Une meilleure manière d'améliorer sensiblement les vitesses de convergence en loi est d'utiliser la méthode d'extrapolation de Romberg.

Talay and Tubaro [57] et Bally et Talay [6, 7] ont montré que l'erreur faible de discrétisation, du schéma d'Euler en particulier, peut s'écrire en développement limité en fraction $\frac{1}{n}$.

Théorème 2.3. On suppose que b et σ sont deux fonctions de classe C^∞ à dérivées bornées. On suppose de plus que

- Soit f est une fonction de classe C^∞ à dérivées à croissance polynômiales.
- Soit f est une fonction mesurable bornée, et l'opérateur σ satisfait une condition d'ellipticité :

$$\exists \epsilon \text{ tel que } \forall x \in \mathbb{R}^d, \|\sigma(x)\sigma^*(x)\| > \epsilon.$$

Alors, pour tous $h = \frac{T}{n}$ l'erreur en T s'écrit

$$\mathbb{E}f(X_T) - \mathbb{E}f\left(\overline{X}_T^h\right) = C(f)h + \mathcal{O}(h^2).$$

où $C(f)$ est une constante qui peut être calculée en fonction de f (voir [6, 7] pour une expression de $C(f)$).

Nous sommes maintenant capable d'appliquer l'extrapolation de Romberg.

Corollaire 2.4. *On suppose les mêmes hypothèses que pour le théorème 2.3. Soit $\bar{X}^{n/2}$ le schéma d'Euler avec pas de $n/2$. Alors*

$$|\mathbb{E}f(X_T) - \{2\mathbb{E}f(\bar{X}_T^h) - \mathbb{E}f(\bar{X}_T^{2h})\}| \leq K_T h^2 .$$

Ce résultat est une conséquence immédiate de 2.3. Le coût numérique de cette méthode est bien plus faible que celui d'un schéma d'ordre 2 (voir [57]).

2.3 Méthodes spéciales pour les options exotiques

Quand le payoff d'une option est spécifié on peut construire des méthodes plus efficaces.

2.3.1 Options asiatiques

Dans ce paragraphe, on suppose que le payoff s'écrit

$$f\left(S_T, \int_0^T S_s ds\right),$$

où f est une fonction bornée et $(S_t, t \geq 0)$ est la solution de l'EDS de Black et Scholes

$$S_t = x e^{(r - \frac{1}{2}\sigma^2)t + \sigma W_t}.$$

Si on veut utiliser des méthodes de Monte Carlo pour calculer le prix d'une de ces options, on doit simuler la moyenne de S_t , et donc approcher son intégrale. Ici, il n'est pas nécessaire d'approcher S_t car il peut être simulé exactement aux instants kh avec $h = T/N$, et on notera alors les instants $t_k = kT/N = kh$. Nous introduisons trois schémas pour approcher $Y_T = \int_0^T S_u du$ (voir [41, 58]).

Le schéma standard

Comme il est facile de simuler S_t à l'instant t , l'intégrale peut être approchée par une somme de Rieman

$$\mathbf{Y}_T^{\mathbf{r}, \mathbf{N}} = \mathbf{h} \sum_{\mathbf{k}=0}^{\mathbf{N}-1} \mathbf{S}_{t_{\mathbf{k}}}. \quad (2.8)$$

Si M représente le nombre de simulation de Monte Carlo, une approximation du prix d'un call fixe asiatique est donnée par

$$\frac{e^{-rT}}{M} \sum_{j=1}^M \left(\frac{h}{T} \sum_{k=0}^{N-1} S_{t_k} - K \right)_+ .$$

La complexité de cet algorithme est en $O\left(\frac{1}{NM}\right)$ (cela est vrai quelque soit l'algorithme Monte Carlo considéré) et il comporte deux types d'erreurs : l'erreur de simulation quantifié par l'écart type et l'erreur due au schéma de discrétisation en h .

Ce schéma (2.8) peut être interprété à l'aide d'une approximation d'Euler de l'équation différentielle stochastique bi-dimensionnelle suivante

$$dU_t = B(U_t)dt + \Sigma(U_t)dW_t \quad \text{avec } U_t = \begin{bmatrix} S_t \\ Y_t \end{bmatrix}, \quad B(U_t) = \begin{bmatrix} rS_t \\ S_t \end{bmatrix} \quad \text{and } \Sigma(U_t) = \begin{bmatrix} \sigma(S_t) \\ 0 \end{bmatrix}$$

Schémas d'ordre supérieur

Une manière de construire des schémas plus précis est de remarquer que dans L^2 , la variable aléatoire "la plus proche" de $(\frac{1}{T} \int_0^T S_s ds - K)_+$, quand les $(S_{t_k}, k = 0, \dots, N)$ sont connus est donnée par

$$\mathbb{E} \left(\left(\frac{1}{T} \int_0^T S_u du - K \right)_+ \middle| \mathcal{B}_h \right), \quad (2.9)$$

où \mathcal{B}_h est la tribu engendrée par les $(S_{t_k}, k = 0, \dots, N)$. Bien évidemment, il est impossible de calculer exactement cette espérance conditionnelle (cela est plus difficile que de calculer une formule explicite pour V).

Mais comme la loi conditionnelle de W_u par rapport à \mathcal{B}_h pour $u \in [t_k, t_{k+1}]$ peut être formellement décrite, on peut calculer

$$\left(\mathbb{E} \left(\frac{1}{T} \int_0^T S_u du \middle| \mathcal{B}_h \right) - K \right)_+ = \left(\frac{1}{T} \int_0^T \mathbb{E} \left(S_u \middle| \mathcal{B}_h \right) du - K \right)_+ \quad (2.10)$$

en tant que fonction de $(W_{t_k}, k = 0, \dots, N)$. L'inégalité de Jensen prouve que (2.10) est plus petit que (2.9), mais on va voir que (2.9) est déjà une bonne approximation de Y_T . Utilisant la loi décrite par

$$\mathcal{L}(W_u \mid W_{t_k} = x, W_{t_{k+1}} = y) = \mathcal{N} \left(\frac{t_{k+1} - u}{h} x + \frac{u - t_k}{h} y, \frac{(t_{k+1} - u)(u - t_k)}{h} \right), \quad (2.11)$$

on obtient

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{T} \int_0^T S_u du \middle| \mathcal{B}_h \right] \\ &= \frac{1}{T} \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} e^{(r - \frac{\sigma^2}{2})u} e^{\sigma \frac{t_{k+1} - u}{h} W_{t_k} + \sigma \frac{u - t_k}{h} W_{t_{k+1}} + \frac{\sigma^2}{2} \frac{(t_{k+1} - u)(u - t_k)}{h}} du \\ &= \frac{1}{T} \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} e^{\sigma \frac{u - t_k}{h} (W_{t_{k+1}} - W_{t_k}) - \frac{\sigma^2}{2} \frac{(u - t_k)^2}{h} + ru} e^{\sigma W_{t_k} - \frac{\sigma^2}{2} t_k} du \end{aligned}$$

Dans une méthode de Monte Carlo, cette approximation est utilisée dans une double boucle (en temps et en nombre de simulations). Il est nécessaire de simplifier cette formule, et, une application de la formule de Taylor (avec h petit) conduit au schéma d'approximation plus simple suivant

$$\mathbf{Y}_T^{\text{e},N} = \frac{h}{T} \sum_{k=0}^{N-1} \mathbf{S}_{t_k} \left(\mathbf{1} + \frac{r h}{2} + \sigma \frac{\mathbf{W}_{t_{k+1}} - \mathbf{W}_{t_k}}{2} \right). \quad (2.12)$$

Remarque 2.6. Notons que ce schéma est "équivalent" à la formule des trapèzes. En effet on montre que

$$\mathbb{E} \left(Y_T^{e,N} - \frac{1}{T} \sum_{k=0}^{N-1} h \frac{S_{t_k} + S_{t_{k+1}}}{2} \right)^2 = O \left(\frac{1}{N^3} \right).$$

Or, comme la vitesse de convergence de (2.12) est en $1/N$.

Démonstration. Ce résultat peut être obtenu en utilisant la formule de Taylor

$$\begin{aligned} \frac{1}{T} \sum_{k=0}^{N-1} h \frac{S_{t_k} + S_{t_{k+1}}}{2} &= \frac{1}{T} \sum_{k=0}^{N-1} \frac{h S_{t_k}}{2} (e^{\sigma(W_{t_{k+1}} - W_{t_k}) - \frac{\sigma^2}{2} h + r h} + 1) \\ &= \frac{1}{T} \sum_{k=0}^{N-1} \frac{h S_{t_k}}{2} (2 + \sigma(W_{t_{k+1}} - W_{t_k}) + r h + O(h(W_{t_{k+1}} - W_{t_k}))) \end{aligned}$$

ce qui est exactement le schéma (2.12). Le terme d'ordre $\sigma^2 h$ et la variation quadratique $\sigma(W_{t_{k+1}} - W_{t_k})$ s'annule. Utilisant Cauchy-Schwarz on obtient le résultat voulu. \square

Le dernier schéma est très semblable. Comme le mouvement Brownien est un processus gaussien, $\int_0^T W_u du$ suit une loi normale sur \mathbb{R} et peut être facilement simulé. Il est alors naturel de chercher des schémas de discrétisation de $\int_0^T S_s ds$ faisant intervenir $\int_0^T W_s ds$. Par exemple, on peut procéder de la manière suivante :

$$\begin{aligned} Y_T &= \frac{1}{T} \int_0^T S_u du \\ &= \frac{1}{T} \sum_{k=0}^{N-1} S_{t_k} \int_{t_k}^{t_{k+1}} e^{\sigma(W_u - W_{t_k}) - \frac{\sigma^2}{2}(u - t_k) + r(u - t_k)} du. \end{aligned}$$

En utilisant formellement la formule de Taylor, on obtient

$$\mathbf{Y}_T^{\mathbf{P},N} = \frac{1}{T} \sum_{k=0}^{N-1} \mathbf{S}_{t_k} \left(\mathbf{h} + \frac{\mathbf{r}h^2}{2} + \sigma \int_{t_k}^{t_{k+1}} (\mathbf{W}_u - \mathbf{W}_{t_k}) d\mathbf{u} \right). \quad (2.13)$$

Remarque 2.7. En pratique pour simuler ce schéma, on doit, à chaque étape simuler $W_{t_{k+1}}$ sachant W_{t_k} et $(\int_{t_k}^{t_{k+1}} W_u du \mid W_{t_k}, W_{t_{k+1}})$. Pour la seconde variable on utilise la loi (2.9) et pour la première on remarque que $(W_{t_{k+1}} - W_{t_k}, k = 0, \dots, N-1)$ est une suite de variables gaussiennes i.i.d.

Remarque 2.8. Ce schéma est généralisable à une classe plus large de processus. Soit S_t une diffusion de dérive $b(S_t)$, de dérive $\sigma(S_t)$, et $Y_t = \int S_t dt$. On peut utiliser un schéma classique pour approcher S_t par S_t^N et poser

$$Y_T^N = \sum_{k=0}^{N-1} S_{t_k}^n \left(h + \int_{t_k}^{t_{k+1}} (S_s^N - S_{t_k}^N) ds \right).$$

Mais si la diffusion S_t est directement simulable aux instants t_k , il suffit d'utiliser

$$Y_T^N = \sum_{k=0}^{N-1} S_{t_k} \left(h + \int_{t_k}^{t_{k+1}} (S_s^N - S_{t_k}) ds \right).$$

Toutefois, si l'ordre de convergence du schéma utilisé pour S n'est pas meilleur que $1/N^{3/2}$, cela n'a pas d'intérêt car l'erreur globale sera limitée par celle du premier schéma.

Convergence dans les espaces L^p

Dans tout ce qui suit on va se ramener à $S_0 = 1$. Si $S_0 = s$ n'est pas une variable aléatoire, il suffit de considérer S_t/s (on note que si $s = 0$ le problème posé est trivial).

Soit un schéma d'approximation de Y_t . Si on s'intéresse à des schémas \mathcal{B}_h mesurables alors on sait que l'espérance conditionnelle va être optimale dans L^2 . Le but est de comparer les différents schémas. On commence par rappeler deux résultats importants.

Proposition 2.5. *Pour une diffusion de type Black et Scholes*

$$\mathbb{E}|S_t - S_s|^{2q} \leq C_q |t - s|^q.$$

Cette proposition est valable pour toute diffusion avec des coefficients lipschitziens (cf. [52]). Le résultat suivant est aussi très utile (voir [23] chapitre 3 pour une preuve) :

Lemme 2.6. *Soit $Z_t = Z_0 + \int_0^t A_s dW_s + \int_0^t B_s ds$ où B_s est un vecteur dans \mathbb{R}^n , A_s une matrice de $\mathbb{R}^{n \times d}$, et W_t un mouvement Brownien d -dimensionnel. (Z_t est un processus d'Itô donc A et B sont adaptés, $\int |A_s| ds < +\infty$ and $\mathbb{E} \int B_s^2 ds < +\infty$).*

Alors, Z_t satisfait

$$\mathbb{E}|Z_t|^p \leq \mathbb{E}|Z_0|^p + C \int_0^t \mathbb{E}(|Z_s|^p + |A_s|^p + |B_s|^p) ds$$

On peut maintenant obtenir des résultats précis de convergence pour nos trois schémas

Proposition 2.7. *Avec les notations ci-dessus, il existe trois fonctions strictement croissantes $K_1(T)$, $K_2(T)$, $K_3(T)$ telles que,*

$$\left(\mathbb{E} \left(\sup_{t \in [0, T]} |Y_t^{r, N} - Y_t|^{2q} \right) \right)^{\frac{1}{2q}} \leq \frac{K_1(T)}{N}, \quad (2.14)$$

$$\left(\mathbb{E} \left(\sup_{t \in [0, T]} |Y_t^{e, N} - Y_t|^{2q} \right) \right)^{\frac{1}{2q}} \leq \frac{K_2(T)}{N}, \quad (2.15)$$

$$\left(\mathbb{E} \left(\sup_{t \in [0, T]} |Y_t^{p, N} - Y_t|^{2q} \right) \right)^{\frac{1}{2q}} \leq \frac{K_3(T)}{N^{3/2}}. \quad (2.16)$$

FIG. 2.1 – Limites de la méthode naïve d'approximation

2.3.2 Options sur maximum

Introduction

On considère dans cette partie des payoff du type

$$f(X_T, M_T),$$

où $(X_t, t \geq 0)$ est solution de l'EDS en dimension 1,

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t,$$

et $M_T = \max_{s \leq T} X_s$.

Approche naïve La méthode d'évaluation la plus simple est de calculer le maximum sur la trajectoire. On va approcher M_T par

$$\bar{M}_T = \max_{0 \leq k \leq n} \bar{X}_{kh}^n,$$

Notons que sous certaines conditions de régularité on montre que (voir [54]) l'erreur s'écrit comme

$$\mathbb{E}(f(X_T, M_T)) - \mathbb{E}(f(\bar{X}_T^n, \bar{M}_T^n)) = \frac{C}{\sqrt{n}}(1 + \epsilon(n)),$$

où $\lim_{n \rightarrow +\infty} \epsilon(n) = 0$. Ce résultat est surprenant. En effet, si f ne dépend pas de M_T , on a vu que l'ordre de convergence était en h . La figure 2.3.2 montre qu'en fait le maximum peut ne pas se situer en un instant de discrétisation et être même très différent. Cependant on peut très facilement améliorer ce résultat (voir [30, 31, 32], [1, 3, 2]).

Utilisation des ponts browniens

L'idée de base est que l'on peut simuler, après discrétisation, la loi du maximum du processus $(\bar{X}_t^n, 0 \leq t \leq T)$ où

$$\bar{X}_t^n = \bar{X}_{kh}^n + b(\bar{X}_{kh}^n)(t - kh) + \sigma(\bar{X}_{kh}^n)(W_t - W_{kh}).$$

conditionnellement à $(\bar{X}_{kh}^n, 0 \leq k \leq N)$. Pour cela, nous avons besoin de quelques notions sur les ponts Browniens :

Proposition 2.8. *Soit $(W_t)_{t \geq 0}$ un mouvement Brownien. Le processus $(Z_t)_{0 \leq t < T}$ défini par $Z_t = W_t - \frac{t}{T}W_T$ est un processus gaussien indépendant de W_T . De plus,*

$$\begin{aligned} \mathbb{E}[Z_t] &= 0 \quad \forall t \in [0, T] \\ \mathbb{E}[Z_t Z_s] &= s \wedge t - \frac{st}{T} \quad \forall (s, t) \in [0, T]^2 \end{aligned}$$

Enfin, le mouvement Brownien conditionné par $W_T = y$ a même loi que

$$Z_t^y = W_t - \frac{t}{T}(W_T - y).$$

Démonstration. Pour tous $0 \leq t_1 < \dots < t_k < T$, le vecteur aléatoire $(Z_{t_1}, \dots, Z_{t_k}, W_T)$ est gaussien centré en tant qu'image d'un vecteur gaussien centré par une application linéaire. En particulier, $(Z_t)_{t \in [0, T]}$ est un processus gaussien centré.

De plus,

$$\text{Cov}(Z_{t_i}, W_T) = \mathbb{E}[W_{t_i} W_T] - \frac{t_i}{T} \mathbb{E}[W_T^2] = 0.$$

d'où l'indépendance de $(Z_{t_1}, \dots, Z_{t_k})$ et W_T puis celle de $Z_t, t \in [0, T]$ et W_T par un argument de classe monotone.

On a déjà vu que Z est centré et pour $s \leq t$

$$\mathbb{E}[Z_s Z_t] = \mathbb{E}[W_s W_t] - \frac{t}{T} \mathbb{E}[W_s W_T] - \frac{s}{T} \mathbb{E}[W_t W_T] + \frac{st}{T^2} \mathbb{E}[W_T^2] = s - \frac{st}{T}.$$

D'où en symétrisant

$$\mathbb{E}[Z_s Z_t] = s \wedge t - \frac{st}{T}.$$

Pour la dernière affirmation, nous renvoyons à [52] p.39. □

Appliquons maintenant ce résultat au schéma d'Euler :

Proposition 2.9. *On suppose que σ ne s'annule pas pour tout $x \in \mathbb{R}$. Alors conditionnellement à $(\bar{X}_{t_k}^n = x_k, \bar{X}_{t_{k+1}}^n = x_{k+1})$, le processus $(\bar{X}^n)_{t \in [t_k, t_{k+1}]}$ a la loi de $(x_k + \sigma(x_k)Z_{t-t_k})$ où Z est le pont Brownien*

$$Z_{t-t_k} = W_{t-t_k} - \frac{t-t_k}{t_{k+1}-t_k} \left(W_{t_{k+1}-t_k} - \frac{x_{k+1}-x_k}{\sigma(x_k)} \right).$$

C'est un processus gaussien d'espérance $x_k \frac{t_{k+1}-t}{t_{k+1}-t_k} + x_{k+1} \frac{t-t_k}{t_{k+1}-t_k}$ et de variance

$$\frac{(t-t_k)(t_{k+1}-t)}{t_{k+1}-t_k} \sigma(x_k)^2.$$

Démonstration. Immédiat d'après la proposition précédente. □

Finalement il est nécessaire d'avoir des renseignements sur la fonction de répartition du maximum d'un pont Brownien :

Proposition 2.10. *Soit $Z_t = W_t - \frac{t}{T}(W_T - y)$, le pont Brownien valant y quand $t = T$. Alors, pour tout $a > y$,*

$$\mathbb{P} \left[\max_{t \in [0, h]} Z_t \leq a \right] = 1 - e^{-\frac{2}{h}a(a-y)}.$$

Démonstration. D'après la proposition 2.8, Z_t a même loi que $W_t \mid W_h = y$. Soit $\tau_a := \inf\{t \geq 0 : W_t = a\}$, le temps d'atteinte de a par W . On a alors,

$$\mathbb{P}\left(\max_{t \in [0, h]} W_t \geq a, W_h \leq y\right) = \mathbb{P}(\tau_a \leq h, W_h \leq y) = \mathbb{P}(\tau_a \leq h, W_h - W_{\tau_a} \leq y - a).$$

Comme τ_a est $\mathcal{F}_{\tau_a}^W$ -mesurable et que $W_h - W_{\tau_a}$ est indépendant de $\mathcal{F}_{\tau_a}^W$ par la propriété de Markov forte du mouvement Brownien, on obtient

$$\begin{aligned} \mathbb{P}\left(\max_{t \in [0, h]} W_t \geq a, W_h \leq b\right) &= \mathbb{P}(\tau_a \leq h, W_h - W_{\tau_a} \geq a - y) \\ &= \mathbb{P}(\tau_a \leq h, W_h \geq 2a - y) \end{aligned}$$

car $W_h - W_{\tau_a}$ et $W_{\tau_a} - W_h$ ont la même loi (propriété de symétrie). Comme $2a - y \geq a$, on a donc

$$\mathbb{P}\left(\max_{t \in [0, h]} W_t \geq a, W_h \leq b\right) = \mathbb{P}(W_h \geq 2a - y).$$

Finalement, comme

$$\mathbb{P}\left(\max_{t \in [0, h]} W_t \leq a \mid W_h = y\right) = 1 - \frac{\frac{\partial}{\partial y} \mathbb{P}(\max_{t \in [0, h]} W_t \geq a, W_h \leq y)}{\frac{\partial}{\partial y} \mathbb{P}(W_h \leq y)},$$

un calcul direct donne le résultat. \square

En couplant les propositions 2.10 et 2.9, on s'aperçoit que la loi du maximum du schéma d'Euler entre t_k et t_{k+1} a pour fonction de répartition :

$$\mathbb{P}\left[\max_{t_k \leq t \leq t_{k+1}} \bar{X}_t^n \leq a \mid \bar{X}_{t_k}^n = x_k, \bar{X}_{t_{k+1}}^n = x_{k+1}\right] = 1 - e^{-\frac{2}{h} \frac{(a-x_k)(a-x_{k+1})}{\sigma^2(x_k)}} := F_h(a, x_k, x_{k+1}).$$

Son inverse étant

$$F_h^{-1}(U, x_k, x_{t_{k+1}}) = \frac{1}{2} \left(x_k + x_{k+1} + \sqrt{(x_{k+1} - x_k)^2 - 2\sigma^2(x_k)h \ln(U)} \right) \quad (2.17)$$

On peut simuler la loi de $\hat{m}_k = \max_{t_k \leq t \leq t_{k+1}} \bar{X}_t^n$ par $F_h^{-1}(U_k, x_k, x_{t_{k+1}})$ où les U_k sont des variables indépendantes de loi uniforme sur $[0, 1]$. On définit une approximation de M_T en utilisant

$$\widehat{M}_T = \sup_{1 \leq k \leq N} \hat{m}_k.$$

Ce schéma est facile à implémenter :

For i=1 to n

$$S(k+1) = S(k) * \exp((r - \text{sigma}^2/2) h + \text{sigma} * \text{sqrt}(h) * \text{gauss}(k))$$

Génère U

Calcul de mk par (2.17)

end

On peut par ailleurs montrer (voir [32]) que sous certaines conditions de régularité, il est d'ordre $1/n$. On peut généraliser ce procédé quand $d > 1$ en prenant le maximum sur chaque composante de X .

2.3.3 Options barrières

Dans cette partie, X désignera la solution de l'équation différentielle stochastique multi-dimensionnelle :

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d.$$

On s'intéresse à l'évaluation d'espérance de la forme :

$$\mathbb{E}[g(X_T)\mathbf{1}_{\tau > T}] \quad \text{où } \tau := \inf\{t \in [0, T] : X_t \notin D\}, \quad (2.18)$$

où D est un borélien de \mathbb{R}^d . τ représente le temps de sortie de cet ensemble, avec pour convention $\inf \emptyset = +\infty$.

Ces options sont en dimension 1 des cas particuliers des options sur maximum.

Approche naïve

On commence par l'approche la plus simple qui consiste à approcher (2.18) par son équivalent discret

$$\mathbb{E}[\mathbf{1}_{\bar{\tau}^n > T} g(\bar{X}_T^n)] \quad \text{où } \bar{\tau}^n := \inf\{t_i, i \in \{0, \dots, n\} : \bar{X}_{t_i}^n \notin D\} \quad (2.19)$$

est l'équivalent discret de τ . Cette quantité est facilement simulable et on a le résultat de convergence suivant démontré dans [32].

Théorème 2.11. *Si D est borné de frontière ∂D de classe¹ C^3 , b et $\sigma \in C^3$ avec a strictement uniformément elliptique² sur D , alors, pour toute fonction mesurable g bornée qui s'annule sur un voisinage de ∂D , on a*

$$\mathbb{E}[\mathbf{1}_{\bar{\tau}^n > T} g(\bar{X}_T^n)] - \mathbb{E}[\mathbf{1}_{\tau > T} g(X_T)] = O\left(\frac{1}{\sqrt{n}}\right).$$

¹i.e. pour tout $y \in \partial D$, il existe un voisinage $V(y)$ de y et un difféomorphisme φ de $V(y) \mapsto B \subset \mathbb{R}^d$ tels que

- (i) $\varphi(V(y) \cap D) \subset \mathbb{R}_+^d := \{x \in \mathbb{R}^d : x^1 \geq 0\}$
- (ii) $\varphi(V(y) \cap \partial D) \subset \partial \mathbb{R}_+^d$
- (iii) $\varphi \in C^3(V(y))$ et $\varphi^{-1} \in C^3(B)$.

²i.e. $\exists \epsilon$ tel que $\forall x \in D \quad |\sigma(x)\sigma^*(x)| > \epsilon$.

Il s'agit d'un résultat assez négatif dans la mesure où l'on perd la vitesse de convergence faible obtenue pour les options vanilla.

Approche par les ponts de diffusion

Dans cette partie, nous présentons une autre approche qui permet d'améliorer la vitesse de convergence faible du Théorème 2.11. C'est l'équivalent pour les options barrières de ce que l'on a vu au chapitre précédant pour les lookback. Cette fois-ci on approche (2.18) par

$$\mathbb{E} [\mathbf{1}_{\tau^n > T} g(\bar{X}_T^n)] \quad \text{où} \quad \tau^n := \inf\{t \in [0, T] : \bar{X}_t^n \notin D\} \quad (2.20)$$

c'est-à-dire que l'on écrit le problème sur le schéma d'Euler continu. On a alors le résultat de convergence suivant démontré dans [33].

Théorème 2.12. *Si D est un demi-espace, b et $\sigma \in C^5$ avec σ strictement uniformément elliptique sur D , alors, pour toute fonction mesurable g bornée qui s'annule sur un voisinage de ∂D , on a*

$$\mathbb{E} [\mathbf{1}_{\tau^n > T} g(\bar{X}_T^n)] - \mathbb{E} [\mathbf{1}_{\tau > T} g(X_T)] = \frac{C_1}{n} + o\left(\frac{1}{n}\right).$$

On retrouve ainsi une vitesse de convergence faible en $1/n$.

Sur les ponts de diffusions Afin d'implémenter cette approximation, on aura besoin du résultat suivant sur la loi du schéma d'Euler continu conditionné. Cette proposition est l'analogie de 2.9 en dimension $d > 1$.

Lemme 2.13. *On suppose que $\gamma(x) = (\sigma(x)\sigma^*(x))^{1/2}$ est inversible pour tout $x \in \mathbb{R}^d$. Alors, conditionnellement à $(\bar{X}_{t_i}^n = x_i, \bar{X}_{t_{i+1}}^n = x_{i+1})$, le processus $(\bar{X}_t^n)_{t_i \leq t \leq t_{i+1}}$ a la loi de*

$$\left(x_i + \gamma(x_i)\tilde{W}_{t-t_i}\right)_{t_i \leq t \leq t_{i+1}} \quad \text{conditionnellement à} \quad \tilde{W}_{t_{i+1}-t_i} = \gamma(x_i)^{-1}(x_{i+1} - x_i)$$

où \tilde{W} est un mouvement brownien. C'est un processus gaussien d'espérance $x_i \frac{t_{i+1}-t}{t_{i+1}-t_i} + x_{i+1} \frac{t-t_i}{t_{i+1}-t_i}$ et de matrice de variance-covariance $\frac{(s-t_i)(t_{i+1}-t)}{t_{i+1}-t_i} \gamma(x_i)^2$ pour tout $t_i \leq s \leq t \leq t_{i+1}$.

Éléments de preuve. Conditionnellement à $\bar{X}_{t_i}^n = x_i$, $(\bar{X}_t^n)_{t_i \leq t \leq t_{i+1}}$ est un processus de diffusion homogène qui admet pour densité de transition

$$p_h(x, z) = \frac{\exp\left\{-\frac{1}{2h}(z-x-b(x)h)^* (\gamma^{-2}(x_i))(z-x-b(x)h)\right\}}{\sqrt{(2\pi h)^d \det[\gamma^2(x_i)]}}.$$

Pour $t_i \leq s \leq t \leq t_{i+1}$

$$\begin{aligned} \frac{\mathbb{P}\left[\bar{X}_s^n \in dx, \bar{X}_t^n \in dy, \bar{X}_{t_{i+1}}^n \in dx_{i+1} \mid \bar{X}_{t_i}^n = x_i\right]}{dxdydx_{i+1}} &= p_{s-t_i}(x_i, x)p_{t-s}(x, y)p_{t_{i+1}-t}(y, x_{i+1}) \\ \mathbb{P}\left[\bar{X}_s^n \in dx, \bar{X}_{t_{i+1}}^n \in dx_{i+1} \mid \bar{X}_{t_i}^n = x_i\right] &= p_{s-t_i}(x_i, x)p_{t_{i+1}-s}(x, x_{i+1})dxdx_{i+1}. \end{aligned}$$

En divisant le premier terme par le second, on obtient :

$$\mathbb{P} \left[\bar{X}_t^n \in dy \mid \bar{X}_s^n = x, \bar{X}_{t_i}^n = x_i, \bar{X}_{t_{i+1}}^n = x_{i+1} \right] = \frac{p_{t-s}(x, y) p_{t_{i+1}-t}(y, x_{i+1})}{p_{t_{i+1}-s}(x, x_{i+1})} dy.$$

Ceci montre que

$$p_{t_i, x_i}^{t_{i+1}, x_{i+1}}(s, x, t, y) := \frac{p_{t-s}(x, y) p_{t_{i+1}-t}(y, x_{i+1})}{p_{t_{i+1}-s}(x, x_{i+1})}$$

est la densité de transition du processus $(\bar{X}_t^n)_{t_i \leq t \leq t_{i+1}}$ conditionnellement à $(\bar{X}_{t_i}^n = x_i, \bar{X}_{t_{i+1}}^n = x_{i+1})$. Le reste en découle par des calculs directs.

Remarque 2.9. En utilisant la propriété de Markov de \bar{X}^n , on vérifie facilement que les processus $(\bar{X}_t^n)_{t_i \leq t \leq t_{i+1}}$ pour i allant de 0 à $n-1$ sont indépendants conditionnellement à $\{\bar{X}_{t_0}^n, \dots, \bar{X}_{t_n}^n\}$.

Implémentation On peut maintenant décrire la méthode. On commence par simuler le schéma d'Euler $(\bar{X}_{t_i}^n)_{i=1}^n$ et on écrit

$$\mathbb{E} [\mathbf{1}_{\tau^n > T} g(\bar{X}_T^n)] = \mathbb{E} [\mathbb{E} [\mathbf{1}_{\tau^n > T} \mid (\bar{X}_{t_i}^n)_{i=1}^n] g(\bar{X}_T^n)]$$

ce qui signifie qu'il faut calculer $\mathbb{E} \mathbf{1}_{\tau^n > T} \mid (\bar{X}_{t_i}^n)_{i=1}^n$, une fois la trajectoire discrète $(\bar{X}_{t_i}^n)_{i=1}^n$ simulée. Évidemment si un des $\bar{X}_{t_i}^n$ n'est pas dans D , il n'y a rien à faire et le payoff de l'option donne 0. On ne calcule donc cette probabilité que si les $\bar{X}_{t_i}^n$ simulés sont tous dans D .

Pour cela, on va utiliser les résultats de la section précédente. Tout d'abord, la Remarque 2.9 implique que

$$\mathbb{E} [\mathbf{1}_{\tau^n > T} \mid (\bar{X}_{t_i}^n)_{i=1}^n] = \prod_{i=0}^{n-1} \mathbb{P} \left(\forall t \in [t_i, t_{i+1}], \bar{X}_t^n \in D \mid \bar{X}_{t_i}^n, \bar{X}_{t_{i+1}}^n \right).$$

Nous allons maintenant montrer que, dans le cas où D est un demi-espace, le calcul est explicite. On écrit D sous la forme

$$D = \{y \in \mathbb{R}^d : \zeta^*(y - \kappa) > 0\} \quad (2.21)$$

i.e. $\partial \bar{D}$ est l'hyperplan passant par $\kappa \in \mathbb{R}^d$ orthogonal à $\zeta \in \mathbb{R}^d$.

Exemple 2.1. Pour une barrière haute U en dimension 1, on a $D = (-\infty, U)$, ce qui donne $\kappa = U$ et $\zeta = -1$.

On fixe $i \in \{0, \dots, n-1\}$. D'après la proposition 2.10, on a

$$\begin{aligned} & \mathbb{P} \left(\exists t \in [t_i, t_{i+1}], \bar{X}_t^n \notin D \mid \bar{X}_{t_i}^n = x_i, \bar{X}_{t_{i+1}}^n = x_{i+1} \right) \\ &= \mathbb{P} \left(\exists t \in [t_i, t_{i+1}], \zeta^* \gamma(x_i) \tilde{W}_{t-t_i} \leq \zeta^*(\kappa - x_i) \mid \tilde{W}_{t_{i+1}-t_i} = \gamma(x_i)^{-1} (x_{i+1} - x_i) \right). \end{aligned}$$

On choisit maintenant une matrice P orthogonale³ telle que $P^1 \cdot = \frac{1}{\|\gamma(x_i)\zeta\|} \zeta^* \gamma(x_i)$. Le processus $\hat{W}_{\cdot-t_i} = (P\tilde{W}_{t-t_i})_{t \in [t_i, t_{i+1}]}$ a la même loi que \tilde{W} . Par ailleurs

$$\zeta^* \gamma(x_i) \tilde{W}_{t-t_i} = \zeta^* \gamma(x_i) P^* \hat{W}_{t-t_i} = \|\gamma(x_i)\zeta\| \hat{W}_{t-t_i}^1$$

car le vecteur ligne $\zeta^* \gamma(x_i) P^*$ a seulement sa première composante non nulle, égale à $\|\gamma(x_i)\zeta\|$. En ré-injectant ce résultat dans les égalités précédentes, on obtient

$$\begin{aligned} & \mathbb{P} \left(\exists t \in [t_i, t_{i+1}], \bar{X}_t^n \notin D \mid \bar{X}_{t_i}^n = x_i, \bar{X}_{t_{i+1}}^n = x_{i+1} \right) \\ &= \mathbb{P} \left(\exists t \in [t_i, t_{i+1}], \hat{W}_{t-t_i}^1 \leq \frac{\zeta^*(\kappa - x_i)}{\|\gamma(x_i)\zeta\|} \mid \hat{W}_{t_{i+1}-t_i}^1 = \frac{\zeta^*(x_{i+1} - x_i)}{\|\gamma(x_i)\zeta\|} \right) \\ &= \mathbb{P} \left(\min_{t \in [t_i, t_{i+1}]} \hat{W}_{t-t_i}^1 \leq \frac{\zeta^*(\kappa - x_i)}{\|\gamma(x_i)\zeta\|} \mid \hat{W}_{t_{i+1}-t_i}^1 = \frac{\zeta^*(x_{i+1} - x_i)}{\|\gamma(x_i)\zeta\|} \right). \end{aligned}$$

Cette probabilité se calcule facilement en utilisant le principe de réflexion du Brownien (voir proposition 2.10)

$$\mathbb{P} \left(\min_{t \in [t_i, t_{i+1}]} \hat{W}_{t-t_i}^1 \leq a \mid \hat{W}_{t_{i+1}-t_i}^1 = b \right) = e^{-2\frac{n}{T}a(a-b)} \quad \forall a \leq 0 \text{ et } b \geq a \quad (2.22)$$

que l'on applique à $b = \frac{\zeta^*(x_{i+1}-x_i)}{\|\gamma(x_i)\zeta\|}$ et $a = \frac{\zeta^*(\kappa-x_i)}{\|\gamma(x_i)\zeta\|}$. Puisque l'on ne fait ce calcul que si $x_i \in D$, on vérifie bien que $a \leq 0$ et $b \geq a$. Finalement, on a montré que

Proposition 2.14. *Si $\gamma(x)^2$ est inversible pour tout $x \in \mathbb{R}^d$ et si D est donné par (2.21), alors pour tout $x_i, x_{i+1} \in D$, on a*

$$\begin{aligned} & \mathbb{P} \left(\forall t \in [t_i, t_{i+1}], \bar{X}_t^n \in D \mid \bar{X}_{t_i}^n = x_i, \bar{X}_{t_{i+1}}^n = x_{i+1} \right) \\ &= 1 - \exp \left(-2\frac{n}{T} \frac{(\zeta^*(\kappa - x_i))^* (\zeta^*(\kappa - x_{i+1}))}{\|\gamma(x_i)\zeta\|^2} \right). \end{aligned} \quad (2.23)$$

Lorsque D n'est pas un demi-espace, on n'a en général plus de forme explicite pour (2.23). On peut toutefois essayer d'approcher D par son hyperplan tangent en $\Pi_{\partial D}(x_i)$, le projeté de x_i sur la frontière de D . Si D est de classe C^5 , on retrouve la vitesse en $1/n$ du Théorème 2.12, voir [33].

Exemple 2.2. *On évalue un call up-and-out*

$$\mathbb{E} \left[\mathbf{1}_{\max_{t \in [0, T]} X_t < U} [X_T - K]^+ \right]$$

dans le modèle de Black-Scholes de paramètres $r = 0$, $\sigma = 0.15$, $T = 1$, $X_0 = 100$, $K = 90$, $U = 130$. Le tableau ci-dessous donne les intervalles de confiance simulés pour les deux méthodes (naïve et par pont), le prix réel étant d'environ 9.21. On effectue à chaque fois 30.000 simulations.

Nombre de pas de temps	Méthode Naïve	Méthode par pont
10	[9.84 , 9.97]	[9.18 , 9.31]
50	[9.46 , 9.60]	[9.14 , 9.27]
100	[9.40 , 9.54]	[9.16 , 9.30]

³i.e. $PP^* = P^*P = I_d$.

La sur-estimation de la méthode naïve est flagrante : il est absolument nécessaire d'utiliser la méthode tenant compte de la probabilité de sortie de D entre deux dates de discrétisation.

Chapitre 3

Réduction de variance

Nous avons vu que l'erreur due à une méthode de Monte Carlo pour le calcul de $\mathbb{E}[f(S_T)]$ est liée soit à la discrétisation du processus S_T soit à l'approximation de l'espérance par une moyenne trajectorielle. Nous avons également observé que dans certains cas de payoff, il était possible de réduire considérablement l'erreur de discrétisation. Dans ce chapitre nous allons nous intéresser au deuxième type d'erreur.

3.1 Fonctions d'importance

3.1.1 Un exemple en finance

Supposons que le processus S obéit à l'équation de Black et Scholes à 1 dimension

$$dS_t = rS_t dt + \sigma S_t dW_t, \quad (3.1)$$

l'on et que l'on veuille calculer un call :

$$\mathbb{E}[\phi(S_T)] = \left(S_0 e^{\sigma W_T + (r - \sigma^2/2)T} - K \right)_+,$$

avec $S_0 \ll K$ (call très en dehors de la monnaie).

Si on effectue une méthode de Monte Carlo classique, sur toutes les trajectoires simulées, très peu seront au dessus de K et donc très peu compteront (car $\mathbb{P} \left(S_0 e^{\sigma \sqrt{T}G + (r - \sigma^2/2)T} > K \right)$ est très faible). Dans ce cas on simule :

$$\mathbb{E}[f(G)] \quad f(x) = \left(S_0 e^{\sigma \sqrt{T}x + (r - \sigma^2/2)T} - K \right)_+.$$

Au lieu de simuler G , nous allons simuler $H = G + m$ où m est réel que l'on choisira ultérieurement. La loi de H a pour densité

$$h(x) = \frac{e^{-\frac{1}{2}(x-m)^2}}{\sqrt{2\pi}}.$$

On a

$$\mathbb{E}[f(G)] = \mathbb{E} \left[f(H) \frac{h(G)}{h(H)} \right] = \mathbb{E} \left[f(H) e^{-mH + \frac{m^2}{2}} \right].$$

Pour choisir un m approprié f doit être donné. Dans le cas du call très en dehors de la monnaie, on va s'arranger pour que $\mathbb{P}[S_0 e^{\sigma\sqrt{T}H+(r-\sigma^2/2)T} > K]$ ne soit plus négligeable. Par exemple, on peut choisir (ce n'est pas optimal) m tel que $S_0 e^{m\sigma\sqrt{T}+(r-\sigma^2/2)T} = K$ et donc

$$\mathbb{P}[S_0 e^{\sigma\sqrt{T}H+(r-\sigma^2/2)T} > K] = \frac{1}{2}.$$

Application : on donne $r = 0$, $\sigma = 0.2$, $T = 1$, $K = 100$, $S_0 = 70$. Le vrai prix est de .248.

Méthode	Nombre de trajectoires	Valeur	Intervalle de confiance
<i>Classique</i>	1000	0.29	[.021 , .57]
<i>Classique</i>	10000	0.288	[.18 , .39]
<i>Classique</i>	100000	0.255	[.23 , .28]
<i>Importance</i>	1000	0.246	[.24 , .25]
<i>Importance</i>	10000	0.239	[.238 , .241]
<i>Importance</i>	100000	0.247	[.247 , .248]

On voit bien qu'un Monte Carlo classique a une variance trop grande.

3.1.2 Généralités

La méthode de fonction d'importance consiste à changer la loi de simulation dans le but de réduire la variance.

Supposons que l'on veuille calculer

$$\mathbb{E}(g(X)),$$

X est une variable aléatoire de densité $f(x)$ on \mathbb{R} , alors

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x)f(x)dx.$$

Soit \tilde{f} une autre densité telle que $\tilde{f}(x) > 0$ et $\int_{\mathbb{R}} \tilde{f}(x)dx = 1$. Alors on peut écrire $\mathbb{E}(g(X))$ comme

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} \frac{g(x)f(x)}{\tilde{f}(x)} \tilde{f}(x)dx = \mathbb{E}\left(\frac{g(Y)f(Y)}{\tilde{f}(Y)}\right),$$

si Y a pour densité $\tilde{f}(x)$ sous \mathbb{P} . Nous avons obtenu une autre méthode pour simuler $\mathbb{E}(g(X))$ en utilisant n trajectoires selon la loi de Y , (Y_1, \dots, Y_n)

$$\frac{1}{n} \left(\frac{g(Y_1)f(Y_1)}{\tilde{f}(Y_1)} + \dots + \frac{g(Y_n)f(Y_n)}{\tilde{f}(Y_n)} \right).$$

On pose $Z = g(Y)f(Y)/\tilde{f}(Y)$, l'algorithme sera alors plus efficace si $\text{Var}(Z) < \text{Var}(g(X))$. Or la variance de Z vaut

$$\text{Var}(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2 = \int_{\mathbb{R}} \frac{g^2(x)f^2(x)}{\tilde{f}(x)} dx - \mathbb{E}(g(X))^2.$$

Si $g(x) > 0$, un calcul simple montre que si $\tilde{f}(x) = g(x)f(x)/\mathbb{E}(g(X))$ alors $\text{Var}(Z) = 0$! Bien sur ce résultat théorique n'a pas d'applications pratiques car il repose sur la connaissance de $\mathbb{E}(g(X))$, qui est ce que l'on veut calculer.

Néanmoins, cela fournit une approche heuristique : on choisit $\tilde{f}(x)$ comme une bonne approximation de $|g(x)f(x)|$. Après normalisation (i.e. on divise par $\int \tilde{f}(x)dx$) on obtient une densité qui peut être simple à simuler.

3.1.3 Théorème de Girsanov et fonctions d'importance pour les diffusions

Théorème Suivant la méthode de Newton [45], nous allons montrer que sous des hypothèses peu contraignantes, la variance peut être annulée.

Proposition 3.1. *Soit Z une variable aléatoire telle que $Z = \psi(W_s, 0 \leq s \leq T)$, $\mathbb{E}(Z^2) < +\infty$ and $\mathbb{P}(Z \geq \epsilon) = 1$, pour un $\epsilon > 0$.*

Il existe un unique processus $(H_t, 0 \leq t \leq T)$ tel que

$$Z = \mathbb{E}(Z) + \int_0^T H_s dW_s.$$

On pose h_t

$$h_t = -\frac{H_t}{\mathbb{E}(Z|\mathcal{F}_t)},$$

Soit

$$L_T = \exp\left(-\int_0^T h_s dW_s - \frac{1}{2}\int_0^T |h_s|^2 ds\right).$$

Alors $\mathbb{E}(L_T) = 1$ et

$$\begin{aligned}\mathbb{E}(Z) &= \tilde{\mathbb{E}}(L_T^{-1}Z) \\ \tilde{\mathbb{P}}(L_T^{-1}Z = \mathbb{E}(Z)) &= 1.\end{aligned}$$

Remarque 3.1. *Cela signifie que sous $\tilde{\mathbb{P}}$, la variable aléatoire $L_T^{-1}Z$ a une variance nulle, i.e. qu'elle est constante p.s.*

Démonstration. Soit $\phi_t = \frac{\mathbb{E}(Z|\mathcal{F}_t)}{\mathbb{E}(Z)}$. On a

$$\phi_t = 1 + \int_0^t \frac{H_s}{\mathbb{E}(Z)} dW_s = 1 - \int_0^t \phi_s h_s dW_s.$$

Cela implique que p.s. sous \mathbb{P} , et sous $\tilde{\mathbb{P}}$

$$\phi_t = \exp\left(-\int_0^t h_s dW_s - \frac{1}{2}\int_0^t |h_s|^2 ds\right) = L_T.$$

Mais, comme Z est une variable aléatoire \mathcal{F}_T -mesurable, $\phi_T = Z/\mathbb{E}(Z)$. \square

Bien entendu, le problème reste difficile car il s'agit maintenant de calculer h . Supposons que le prix du sous-jacent peut être modélisé par une diffusion X_t . Le prix d'une option européenne est alors :

$$\mathbb{E}[e^{-rT} f(X_T)].$$

où f est une fonction continue et positive et $(X_t, t \geq 0)$ une diffusion solution de

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, X_0 = x. \quad (3.2)$$

Soit u la fonction définie par $u(t, x) = \mathbb{E}_x(e^{-r(T-t)} f(X_{T-t}))$ (x est le point de départ de la diffusion). Il est connu que u est solution du problème EDP suivant :

$$\begin{cases} u(T, x) &= f(x), \quad \text{for } x \in \mathbb{R}^n, \\ \left(\frac{\partial u}{\partial t} + Au - ru\right)(t, x) &= 0, \quad \text{for } (t, x) \in [0, T] \times \mathbb{R}^n, \end{cases}$$

où A est le générateur infinitésimal de la diffusion $(X_t, t \geq 0)$.

Si on pose $Z = e^{-rT}u(T, X_T)$ et

$$H_t = e^{-rt} \frac{\partial u}{\partial x}(t, X_t) \sigma(X_t),$$

alors d'après la formule d'Itô, on a

$$e^{-rT}u(T, X_T) = u(0, x) + \int_0^T H_t dW_t = \mathbb{E}[Z] \int_0^T H_t dW_t.$$

Le processus h sera alors défini par :

$$h_t = -\frac{H_t}{\mathbb{E}[Z|\mathcal{F}_t]} = -\frac{\frac{\partial u}{\partial x}(t, X_t) \sigma(X_t)}{u(t, X_t)}.$$

La connaissance d'une approximation de u (par des méthodes liées aux EDP par exemple) permet de calculer h (voir [53] pour des approximations de u par des grandes déviations).

3.2 Variables antithétiques

Le principe de cette méthode de contrôle est d'utiliser des propriétés de symétrie de la loi simulée pour réduire la variance. Dans le cas de la finance on doit souvent calculer $M = \mathbb{E}[\phi(G)]$ où G est une gaussienne centrée. Or on sait que $G \stackrel{\text{loi}}{=} -G$. Donc, un estimateur de $M = \mathbb{E}[\phi(G)]$ est

$$\bar{M}_n = \frac{1}{2n}(\phi(G_1) + \phi(-G_1) + \dots + \phi(G_n) + \phi(-G_n)).$$

où G_1, \dots, G_n sont n réalisations de la loi de G . Si on note $M_n = \frac{1}{n}(\phi(G_1) + \dots + \phi(G_n))$ l'estimateur Monte Carlo classique, on obtient

$$\begin{aligned} \text{Var}(M_{2n}) &= \frac{1}{4n^2} \text{Var} \left(\sum_{i=1}^{2n} \phi(G_i) \right) \stackrel{G_i \text{ indépendantes}}{=} \frac{1}{2n} \sum_{i=1}^{2n} \text{Var}(\phi(G_i)) \\ &\stackrel{G_i \text{ de même loi}}{=} \frac{1}{2n} \text{Var}(\phi(G_1)). \end{aligned}$$

La variance de l'estimateur \bar{M} est

$$\begin{aligned} \text{Var}(\bar{M}_{2n}) &= \frac{1}{4n^2} \text{Var} \left(\sum_{i=1}^n (\phi(G_i) + \phi(-G_i)) \right) \\ &\stackrel{G_i \text{ indépendantes}}{=} \frac{1}{4n^2} \sum_{i=1}^n \text{Var}(\phi(G_i) + \phi(-G_i)) \\ &\stackrel{G_i \text{ de même loi}}{=} \frac{1}{4n} (\text{Var}(\phi(G_1)) + \text{Var}(\phi(-G_1)) + 2 * \text{Cov}(\phi(G_1), \phi(-G_1))) \\ &\stackrel{G_1 \stackrel{\text{loi}}{=} -G_1}{=} \frac{1}{2n} (\text{Var}(\phi(G_1)) + \text{Cov}(\phi(G_1), \phi(-G_1))) \end{aligned}$$

On conclut que $\text{Var}(\bar{M}_{2n}) \leq \text{Var}(M_{2n})$ si et seulement si $\text{Cov}(\phi(G_1), \phi(-G_1)) \leq 0$. Or, on dispose du théorème :

Théorème 3.2. Soit G une variable aléatoire, T une transformation décroissante de \mathbb{R} telle que $T(G) \stackrel{\text{loi}}{=} G$, et ϕ une fonction monotone alors

$$\text{Cov}(\phi(G), \phi(T(G))) \leq 0,$$

avec inégalité si ϕ est strictement monotone sur un domaine de mesure non nulle.

Démonstration. Nous ferons la preuve dans le cas où ϕ est croissante et T est décroissante (les autres cas s'obtiennent par symétrie). Soit H une variable aléatoire indépendante de G et de même loi, alors

$$\begin{aligned} \mathbb{E}[\phi(G)] \mathbb{E}[\phi(T(G))] &= \mathbb{E}[\phi(G)\phi(T(H))] = \mathbb{E}[(\phi(G) - \phi(H))(\phi(T(H)) - \phi(T(G)))] \\ &\quad + \mathbb{E}[\phi(H)(\phi(T(H)) - \phi(T(G)))] + \mathbb{E}[\phi(G)\phi(T(G))] \\ &= \underbrace{\mathbb{E}[(\phi(G) - \phi(H))(\phi(T(H)) - \phi(T(G)))]}_{(1)} \\ &\quad - \mathbb{E}[\phi(H)\phi(T(G))] + 2\mathbb{E}[\phi(G)\phi(T(G))]. \end{aligned}$$

En regroupant les termes identiques de part et d'autre et en décomposant (1) suivant $\mathbf{1}_{G \geq H}$ on a

$$\begin{aligned} 2\mathbb{E}[\phi(G)] \mathbb{E}[\phi(T(G))] &= \mathbb{E}[(\phi(G) - \phi(H))(\phi(T(H)) - \phi(T(G)))\mathbf{1}_{G \geq H}] \\ &\quad + \mathbb{E}[(\phi(G) - \phi(H))(\phi(T(H)) - \phi(T(G)))\mathbf{1}_{G \leq H}] \\ &\quad + 2\mathbb{E}[\phi(G)\phi(T(G))] \end{aligned}$$

Or comme ϕ est croissante et T décroissante on a

$$(\phi(G) - \phi(H))(\phi(T(H)) - \phi(T(G)))\mathbf{1}_{G \geq H} \geq 0$$

et

$$(\phi(G) - \phi(H))(\phi(T(H)) - \phi(T(G)))\mathbf{1}_{G \leq H} \geq 0.$$

D'où

$$\mathbb{E}[\phi(G)] \mathbb{E}[\phi(T(G))] \geq \mathbb{E}[\phi(G)\phi(T(G))].$$

Il y égalité lorsque $(\phi(G) - \phi(H))(\phi(T(H)) - \phi(T(G)))$ est non nulle sur un ensemble de mesure non nulle, i.e. si ϕ est strictement monotone sur un ensemble de mesure non nulle. \square

Un exemple en finance Considérons le cas du put dans le modèle de Black et Scholes. On cherche à calculer $\mathbb{E}[\phi(G)]$ où $\phi(x) = \left(K - e^{\sigma\sqrt{T}x + \delta T}\right)_+$ et G est une variable aléatoire gaussienne centrée réduite. ϕ est croissante et la transformation $T(x) = -x$ est décroissante. Vu le théorème 3.2, l'estimateur \bar{M}_n a une variance plus faible que M_{2n} . En effet, en prenant $r = 0$, $\sigma = 0.2$, $T = 1$, $K = 100$, $S_0 = 100$, le vrai prix est 7.96.

Méthode	Valeur	Intervalle de confiance
<i>Classique</i>	8.85	[1.49 , 16.21]
<i>Antithétique</i>	8.28	[3.26 , 16.29]

3.3 Variables de contrôles

Supposons que l'on cherche à calculer $\mathbb{E}[X]$ où X est une variable aléatoire donnée. Les méthodes de variables de contrôles consistent à trouver une variable Y d'espérance nulle ou calculable explicitement telle que :

$$\text{Var}(X - Y) \ll \text{Var}(X).$$

La relation $\mathbb{E}[X] = \mathbb{E}[X - Y] + \mathbb{E}[Y]$ permet de calculer $\mathbb{E}[X]$ en simulant $X - Y$. Comme la variance est plus faible, l'estimateur sera de meilleure qualité.

3.3.1 Théorie

On peut montrer que dans le cas où l'on cherche à calculer $\mathbb{E}[\psi(X_t, t \geq 0)]$ où X_t est une diffusion, une variable de contrôle parfaite existe. Ce résultat théorique n'a pas toujours des applications pratiques mais il peut conduire à des procédures efficaces. Le théorème qui permet cela est le théorème de représentation prévisible :

Théorème 3.3. *Soit Z une variable aléatoire telle que $\mathbb{E}(Z^2) < +\infty$. Supposons que Z est mesurable par rapport à la σ -algèbre engendrée par $\sigma(W_s, s \leq T)$. Alors, il existe un processus stochastique $(H_t, t \leq T)$ adapté à $\sigma(W_s, s \leq t)$, tel que $\mathbb{E}\left(\int_0^T H_s^2 ds\right) < +\infty$ et*

$$Z = \mathbb{E}(Z) + \int_0^T H_s dW_s.$$

On peut trouver une preuve dans ce théorème dans [52] ou [37].

Remarque 3.2. *On peut remarquer que Z doit être mesurable par rapport à la tribu engendrée par le mouvement Brownien.*

Ce théorème prouve qu'en principe, nous sommes capables d'annuler la variance de Z . Mais le calcul explicite de H n'est jamais simple, parfois plus compliqué que celui de $\mathbb{E}[Z]$. Nous renvoyons à [45] pour des approximations numériques et des applications en finance.

Supposons que le prix du sous-jacent soit un modèle Markovien X_t et que le payoff soit une fonction de cette diffusion au temps T . Alors le processus $(H_t, t \leq T)$ peut s'écrire comme $H_t = v(t, X_t)$, v étant une fonction de t et x :

Théorème 3.4. *Soit b et σ deux fonctions lipschitziennes. Soit $(X_t, t \geq 0)$ l'unique solution de*

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, X_0 = x.$$

On note A le générateur infinitésimal de cette diffusion

$$Af(x) = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{j=1}^n b_j(x) \frac{\partial f}{\partial x_j}(x),$$

où $a_{ij}(x) = \sum_{k=1}^p \sigma_{ik}(x)\sigma_{jk}(x)$.

Supposons que u est une fonction $C^{1,2}$ avec des dérivées bornées en x et qu'elle est solution de l'EDP :

$$\begin{cases} \left(\frac{\partial u}{\partial t} + Au \right) (t, x) = f(x), & (t, x) \in [0, T] \times \mathbb{R}^n, \\ u(T, x) = g(x), & x \in \mathbb{R}^n. \end{cases} \quad (3.3)$$

Alors si $Z = g(X_T) - \int_0^T f(X_s) ds$ et $Y = \int_0^T \frac{\partial u}{\partial x}(s, X_s) \sigma(s, X_s) dW_s$, on a

$$\mathbb{E}(Z) = Z - Y.$$

Cela signifie que la variable aléatoire Y est un contrôle parfait de Z .

Démonstration. Grâce à la formule d'Itô, on a

$$du(t, X_t) = \left(\frac{\partial u}{\partial t} + Au \right) (t, X_t) dt + \frac{\partial u}{\partial x}(t, X_t) \sigma(s, X_s) dW_t.$$

Maintenant, en intégrant entre 0 et T , en prenant l'espérance et en utilisant le fait que u soit solution de 3.3, on obtient

$$u(0, x) = Z - Y = \mathbb{E}(Z).$$

□

Remarque 3.3. Évidemment, si l'on pouvait calculer une approximation suffisamment fine de la solution de 3.3, on n'aurait pas besoin d'utiliser de méthode de Monte-Carlo. Toutefois, on peut se restreindre à une résolution grossière de 3.3 qui soit suffisante pour être utilisée dans la méthode de réduction de variance. On peut également approcher le gradient par celui correspondant à un payoff/modèle proche pour lequel on a une formule explicite. Par exemple, dans un modèle à volatilité stochastique, on peut approcher le delta en t_i par le delta de Black-Scholes correspondant à la valeur de la volatilité en t_i .

Remarque 3.4. Cette approche par EDP s'étend aux options asiatiques, lookback ou à barrière. Le cas des options asiatiques est traité simplement en augmentant la taille du processus X , le résultat est similaire à celui du Théorème 1.3 pour le processus augmenté, voir également [40] Proposition 5.2.11. On pourra consulter [54] pour les options lookback et [33] pour les options à barrière.

3.3.2 Exemples

Exemple 3.1. *Parité Call/Put* On suppose que $(S_t, t \geq 0)$ suit une diffusion log-normale

$$dS_t = S_t (r dt + \sigma dW_t), S_0 = x.$$

On note C le prix du Call Européen

$$C = \mathbb{E} \left(e^{-rT} (S_T - K)_+ \right),$$

et P le prix du Put Européen

$$P = \mathbb{E} \left(e^{-rT} (K - S_T)_+ \right).$$

Il existe une relation qui lie le prix du Call et du Put et qui ne dépend pas du modèle. Cette relation peut s'établir par des stratégies d'arbitrage :

$$C - P = \mathbb{E} \left(e^{-rT} (S_T - K) \right) = x - Ke^{-rT}.$$

Cette formule peut également être utilisée pour réduire la variance :

$$C = \mathbb{E} \left(e^{-rT} (K - S_T)_+ \right) + x - Ke^{-rT}.$$

Le calcul du Call est alors déduit du calcul du Put. Or la fonction de Payoff du Put étant bornée, on peut s'attendre à une variance plus faible.

Exemple 3.2. *Méthode de Kemna-Vorst [38] pour les options asiatiques.*

On suppose que $(S_t, t \geq 0)$ suit une diffusion log-normale

$$dS_t = S_t (r dt + \sigma dW_t), S_0 = x.$$

On cherche à calculer le prix d'une option asiatique de payoff

$$\left(\frac{1}{T} \int_0^T S_u du - K \right)_+.$$

On peut simuler le processus moyenne de S_T en utilisant l'un des schémas vus dans le premier chapitre (2.8, 2.12, 2.13).

Pour accroître l'efficacité des simulations, une technique de réduction de variance peut être utilisée. Nous suivons ici la méthode de Kemna et Vorst [38]. Cela consiste à approcher $\frac{1}{T} \int_0^T S_u du$ par $\exp \left(\frac{1}{T} \int_0^T \log(S_u) du \right)$. On peut s'attendre à ce que ces deux variables aléatoires soient proches quand r et σ sont faibles.

La variable aléatoire $Z' = \frac{1}{T} \int_0^T \log(S_u) du$ suit visiblement une loi normale

$$Z' \sim \mathcal{N}(T(r/2 - \sigma^2/4), \sigma^3 T/3).$$

On peut donc calculer explicitement

$$\mathbb{E}(e^{-rT} (\exp(Z') - K)_+).$$

En conséquence, on choisit la variable Z définie par

$$Z = e^{-rT} (x e^{(r - \frac{\sigma^2}{2})\frac{T}{2} + \frac{\sigma}{T} \int_0^T W_u du} - K)_+,$$

comme notre variable de contrôle.

Il faut remarquer que la variable de contrôle doit être calculée avec le chemin du mouvement Brownien déjà simulé. En conséquence, chaque variable de contrôle doit être adaptée au schéma pour laquelle elle est utilisée. On va donc effectuer la même approximation pour $\int_0^T W_u du$ que pour $\int_0^T S_u du$

$$\text{Pour (2.8)} \quad Z_T^{r,N} = e^{-rT} (x e^{(r - \frac{\sigma^2}{2})\frac{T}{2} + \frac{\sigma}{T} \sum_{k=0}^{N-1} h W_{t_k}} - K)_+$$

$$\text{Pour (2.12)} \quad Z_T^{e,n} = e^{-rT} (x e^{(r - \frac{\sigma^2}{2})\frac{T}{2} + \frac{\sigma}{T} \sum_{k=0}^{N-1} \frac{h}{2} (W_{t_k} + W_{t_{k+1}})} - K)_+$$

$$\text{Pour (2.13)} \quad Z_T^{p,N} = e^{-rT} (x e^{(r - \frac{\sigma^2}{2})\frac{T}{2} + \frac{\sigma}{T} \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} W_u du} - K)_+$$

On teste cette méthode pour évaluer un call asiatique de paramètres $r = 0.1$, $\sigma = 0.2$, $T = 1$, $X_0 = K = 100$. On utilise la discrétisation par trapèzes 2.12 pour approcher $\int_0^T X_t dt$ avec $n = 50$. La formule de Black-Scholes permet de calculer $e^{-rT} \mathbb{E}g(Z_T) = 6.77$ et on prend pour variable de contrôle

$$Y = e^{-rT} \left(\mathbb{E}g(Z_T) - \left[e^{(r-\sigma^2/2)\frac{T}{2} + \frac{\sigma}{2n} \sum_{i=0}^{n-1} (W_{t_i} + W_{t_{i+1}})} - K \right]^+ \right).$$

Le tableau ci-dessous donne les intervalles de confiance simulés avec et sans cette variable de contrôle, le prix réel étant d'environ 7.04.

Nombre de simulations	Avec	Sans
10 000	[7.034 , 7.049]	[6.957 , 7.296]
20 000	[7.034 , 7.046]	[6.970 , 7.207]
50 000	[7.039 , 7.045]	[6.974 , 7.124]
100 000	[7.037 , 7.042]	[6.980 , 7.086]

Exemple 3.3. *Options sur Panier*

Une idée similaire peut être utilisée pour les options sur paniers. On suppose que pour $i = 1, \dots, d$

$$S_T^i = x_i e^{rT + \sum_{j=1}^p \sigma_{ij} W_T^j},$$

où W^1, \dots, W^p sont des mouvements Browniens indépendants. Soit a_i , $1 \leq i \leq p$ des réels strictement positifs vérifiant $a_1 + \dots + a_d = 1$. Le but ici est de calculer un Put sur panier :

$$\mathbb{E}((K - X)_+),$$

où $X = a_1 S_T^1 + \dots + a_d S_T^d$. L'idée est d'approcher

$$\frac{X}{m} = \frac{a_1 x_1}{m} e^{rT + \sum_{j=1}^p \sigma_{1j} W_T^j} + \dots + \frac{a_d x_d}{m} e^{rT + \sum_{j=1}^p \sigma_{dj} W_T^j}$$

où $m = a_1 x_1 + \dots + a_d x_d$, par $\frac{Y}{m}$ et Y est une variable aléatoire log-normale définie par

$$Y = m e^{\sum_{i=1}^d \frac{a_i x_i}{m} (rT + \sum_{j=1}^p \sigma_{ij} W_T^j)}.$$

Comme Y suit une loi log-normale, par analogie avec la formule de Black et Scholes, on sait calculer

$$\mathbb{E}[(K - Y)_+].$$

L'utilisation de $Z = (K - Y)_+$ comme variable de contrôle et la simulation de $(K - X)_+ - (K - Y)_+$ fournissent un algorithme efficace.

3.3.3 Valeur moyenne et conditionnement

Cette méthode utilise le fait que le conditionnement réduit la variance. Soit Z une variable aléatoire de carré intégrable, on a

$$\mathbb{E}(Z) = \mathbb{E}(\mathbb{E}(Z|Y)),$$

où Y est n'importe quelle variable aléatoire construite sur le même espace de probabilité. Il est connu que $\mathbb{E}(Z|Y)$ peut s'écrire comme

$$\mathbb{E}(Z|Y) = \phi(Y).$$

De plus, on a $\mathbb{E}(Z) = \mathbb{E}(\phi(Y))$ et comme l'espérance conditionnelle est une projection L^2 ,

$$\mathbb{E}(\phi(Y)^2) \leq \mathbb{E}(Z^2),$$

donc $\text{Var}(\phi(Y)) \leq \text{Var}(Z)$.

Bien sur, l'efficacité de la méthode repose sur la connaissance de la fonction ϕ . Cela est clair lorsque $Z = f(X, Y)$, où X et Y sont des variables indépendantes. Dans ce cas,

$$\mathbb{E}(f(X, Y)|Y) = \phi(Y),$$

où $\phi(y) = \mathbb{E}(f(X, y))$.

Exemple 3.4. *Modèle à volatilité stochastique*

Soit $(W_t, t \geq 0)$ un mouvement Brownien. On suppose que $(S_t, t \geq 0)$ suit un modèle à volatilité stochastique défini par

$$dS_t = S_t (rdt + \sigma_t dW_t^1), S_0 = x,$$

où $(\sigma_t, t \geq 0)$ est un processus stochastique indépendant de $(W_t, t \geq 0)$. Nous voulons calculer, en utilisant des techniques de Monte Carlo

$$\mathbb{E}(e^{-rT} f(S_T)),$$

où f est une fonction de payoff. Clairement S_T s'exprime comme

$$S_T = x \exp \left(rT - \int_0^T \sigma_t^2 / 2 dt + \int_0^T \sigma_t dW_t^1 \right).$$

Mais, comme les processus $(\sigma_t, t \geq 0)$ et $(W_t, t \geq 0)$ sont indépendants,

$$\int_0^T \sigma_t dW_t \text{ est égal en loi à } \sqrt{\frac{1}{T} \int_0^T \sigma_t^2 dt} \times W_T.$$

En conditionnant par rapport au processus σ on obtient

$$\mathbb{E}(e^{-rT} f(S_T)) = \mathbb{E}(\psi(\sigma_t, 0 \leq t \leq T)),$$

où pour une trajectoire de volatilité fixée ($v_t, 0 \leq t \leq T$)

$$\begin{aligned} \psi(v_t, 0 \leq t \leq T) &= \mathbb{E} \left(e^{-rT} f \left(x e^{rT - \int_0^T \frac{v_t^2}{2} dt + \sqrt{\frac{1}{T} \int_0^T v_t^2 dt} \times W_T \right) \right) \\ &= \phi \left(\sqrt{\frac{1}{T} \int_0^T v_t^2 dt} \right). \end{aligned}$$

Or $\phi(\sigma)$ est donné par la formule de Black Scholes

$$\phi(\sigma) = \mathbb{E} \left(e^{-rT} f \left(x \exp \left(\left(r - \frac{\sigma^2}{2} \right) T + \sigma W_T \right) \right) \right).$$

Ainsi quand f est le payoff du call ou du Put, il peut s'exprimer directement en utilisant les résultats de Black Scholes.

Chapitre 4

Calcul des sensibilités

Les sensibilités ou variations du portefeuille par rapport aux paramètres sont importantes en finance. En effet, ce sont ces calculs qui vont permettre au vendeur d'une option de se couvrir. Par ailleurs, nous avons vu que la connaissance du delta permettait de mettre en oeuvre des techniques de réduction de variance.

4.1 Rappels sur les sensibilités

Soit X_T un actif sous jacent uni-dimensionnel solution de

$$dX_t = rX_t dt + \sigma(t, X_t) dW_t,$$

où σ est une fonction. On suppose que l'on cherche à calculer les sensibilités d'une option de payoff $g(X_t, 0 \leq t \leq T)$ par Monte Carlo. Si $X_0 = x$, on note $u(t, x) = \mathbb{E}[g(X_s, 0 \leq s \leq T)]$. Les sensibilités que nous allons aborder ici sont :

$$\begin{aligned}\Delta &= \frac{\partial}{\partial x} \mathbb{E} [e^{-rT} g(X_s, 0 \leq s \leq T)] \\ \Gamma &= \frac{\partial^2}{\partial x^2} \mathbb{E} [e^{-rT} g(X_s, 0 \leq s \leq T)] \\ \mathcal{V} &= \frac{\partial}{\partial \epsilon} \mathbb{E} [e^{-rT} g(X_s^{\sigma(\cdot) + \epsilon \sigma'(\cdot)}, 0 \leq s \leq T)]\end{aligned}$$

Le Δ est la couverture. Il représente la quantité d'actif risqué que l'on doit détenir pour répliquer l'option. Le Γ est la dérivée du Δ et doit rester faible sinon cela signifie que le nombre d'interventions sera trop élevé.

Le \mathcal{V} (Vega) est la dérivé par rapport à la volatilité. La volatilité est le seul paramètre difficile à évaluer. Des méthodes de calibration sont nécessaires. Connaître la sensibilité du prix par rapport à σ permet de jauger la fiabilité de son prix par rapport au processus de calibration. Il existe d'autres sensibilités que nous aborderons peu ou pas dans ce chapitre :

$$\begin{aligned}\Theta &= \frac{\partial}{\partial T} \mathbb{E} [e^{-rT} g(X_s, 0 \leq s \leq T)] \\ \rho &= \frac{\partial}{\partial r} \mathbb{E} [e^{-rT} g(X_s, 0 \leq s \leq T)]\end{aligned}$$

4.2 Approche par différences finies

L'approche la plus simple pour calculer les Grecques (sensibilités) est de voir que, par exemple, $\Delta = \frac{\partial u}{\partial x}(0, x)$. Donc, une méthode simple d'approximation sera :

$$\begin{aligned}\Delta &= \frac{\partial u}{\partial x}(0, x) \sim \frac{1}{2\epsilon} (u(0, x + \epsilon) - u(0, x - \epsilon)) = \tilde{\Delta} \\ \Gamma &= \frac{\partial^2 u}{\partial x^2}(0, x) \sim \frac{1}{2\epsilon} (\tilde{\Delta}(0, x + \epsilon) - \tilde{\Delta}(0, x - \epsilon)) = \tilde{\Gamma}\end{aligned}$$

Pour le \mathcal{V} , la méthode est la même. On génère une première trajectoire avec $\sigma + \epsilon$, une autre avec $\sigma - \epsilon$ et on calcule le taux d'accroissement.

Il existe deux possibilités de simulation de ces accroissements :

1. On utilise M simulations pour estimer une approximation $\hat{u}(0, x + \epsilon)$ de $u(0, x + \epsilon)$ et M autres (indépendantes) pour estimer $\hat{u}(0, x - \epsilon)$ approximation de $u(0, x - \epsilon)$. Dans ce cas, on a :

$$\begin{aligned}\text{Var} \left(\frac{\hat{u}(0, x + \epsilon) - \hat{u}(0, x - \epsilon)}{2\epsilon} \right) &= \frac{1}{4\epsilon^2} (\text{Var}(\hat{u}(0, x + \epsilon)) + \text{Var}(\hat{u}(0, x - \epsilon))) \\ &\sim \frac{1}{4\epsilon^2} \left(\frac{\text{Var}(g(X_T^x))}{N} + \frac{\text{Var}(g(X_T^x))}{N} \right) \\ &= \frac{1}{2N\epsilon^2} \text{Var}(g(X_T^x)).\end{aligned}$$

2. On simule N trajectoires du Brownien et on construit $X^{x+\epsilon}$ et $X^{x-\epsilon}$ avec les mêmes trajectoires. Dans ce cas

$$\begin{aligned}\text{Var} \left(\frac{\hat{u}(0, x + \epsilon) - \hat{u}(0, x - \epsilon)}{2\epsilon} \right) &= \frac{1}{N} \text{Var} \left(\frac{g(X_T^{x+\epsilon}) - g(X_T^{x-\epsilon})}{2\epsilon} \right) \\ &\sim \frac{1}{N} \text{Var}(g'(X_T^x)).\end{aligned}$$

Si ϵ est petit et g régulière, la seconde méthode sera en général préférable à la première. Cette méthode est très simple à mettre en oeuvre mais le choix du ϵ n'est pas évident. Si ϵ est trop petit, la variance de l'estimateur peut être très grande, c'est le cas si le payoff est très irrégulier. Si ϵ est trop grand, l'approximation des dérivées est mauvaise. Les vitesses de convergence de ces méthodes ont été étudiées par [27], [28] et [42].

4.3 Amélioration des techniques pour Monte Carlo

4.3.1 Notion de processus tangent

Nous nous plaçons dans un cadre général. Soit X_t une diffusion d -dimensionnelle solution de l'équation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x,$$

où W est un Brownien r dimensionnel, $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ et $\sigma : \mathbb{R}^d \mapsto \mathbb{R}^{d \times r}$. On étudie la fonction qui à $x \mapsto X_t(x)$, à ω fixé.

Théorème 4.1. *Si b et σ sont deux fonctions \mathcal{C}_b^∞ , alors l'application $x \rightarrow X_t(x)$ est presque sûrement \mathcal{C}^∞ et l'équation d'Itô satisfaite par $\frac{\partial X_t^i(x)}{\partial x^j}$ s'obtient par différentiation sous les intégrales :*

$$\frac{\partial X_t^i(x)}{\partial x^j} = \sum_{k=1}^d \int_0^T \frac{\partial b^i}{\partial x^k}(X_s) \frac{\partial X_s^k(x)}{\partial x^j} ds + \sum_{k=1}^d \sum_{l=1}^r \int_0^T \frac{\partial \sigma^{i,l}}{\partial x^k}(X_s) \frac{\partial X_s^k(x)}{\partial x^j} dW_s^l.$$

que l'on peut écrire :

$$\nabla X_t(x) = \int_0^T \nabla b(X_s(x)) \times \nabla X_s(x) ds + \sum_{l=1}^r \int_0^T \nabla \sigma^l(X_s(x)) \times \nabla X_s(x) dW_s^l,$$

où σ^l est la $l^{\text{ème}}$ colonne de σ .

Remarque 4.1. *On peut différentier le processus X indéfiniment par rapport à x , dès que les fonctions b et σ sont assez régulières.*

Définition 4.1. *Le processus tangent $(Y_t)_{t \geq 0}$ est le processus (matriciel) dérivée première de $(X_t(x))_{t \geq 0}$ par rapport à x*

$$Y_t = \frac{\partial X_t(x)}{\partial x}.$$

En dimension 1, il satisfait l'équation différentielle stochastique :

$$X_t = x + \int_0^T b(X_s) ds + \int_0^T \sigma(X_s) dW_s \quad (4.1)$$

$$Y_t = 1 + \int_0^T b'(X_s) Y_s ds + \int_0^T \sigma'(X_s) Y_s dW_s \quad (4.2)$$

Remarque 4.2. *En dimension 1, l'équation dont le processus tangent est solution a une forme explicite :*

$$Y_t = \exp \left(\int_0^T \left(b'(X_s) - \frac{\sigma'^2}{2}(X_s) \right) ds + \int_0^T \sigma'(X_s) dW_s \right).$$

Démonstration. On considère tout d'abord

$$\begin{aligned} Z_t &= \log \left(\exp \left(\int_0^T \left(b'(X_s) - \frac{\sigma'^2}{2}(X_s) \right) ds + \int_0^T \sigma'(X_s) dW_s \right) \right) \\ &= \int_0^t \left(b'(X_s) - \frac{\sigma'^2}{2}(X_s) \right) ds + \int_0^t \sigma'(X_s) dW_s \end{aligned}$$

Donc $dZ_t = \left(b'(X_t) - \frac{\sigma'^2}{2}(X_t) \right) dt + \sigma'(X_t) dW_t$. De plus, en appliquant la formule d'Itô à $\log(Y_t)$, on a

$$\begin{aligned} d(\log(Y_t)) &= \frac{dY_t}{Y_t} - \frac{1}{2Y_t^2} d\langle Y, Y \rangle_t \\ &= \int_0^t b'(X_s) ds + \int_0^t \sigma'(X_s) dW_s - \int_0^t \frac{\sigma'^2}{2}(X_s) ds = dZ_t \end{aligned}$$

Comme $Z_0 = 0 = \log(Y_0)$ on conclut que $Y_t = \exp(Z_t)$. □

Remarque 4.3. On peut remarquer qu'en dimension $\frac{\partial}{\partial x}X_t(x)$ est une fonction positive et donc que $X_t(x)$ est une fonction croissante en x .

Exemple 4.1. On considère le modèle de Black et Scholes, i.e. $b(x) = rx$ et $\sigma(x) = \sigma x$. Alors on sait que

$$X_t = x \exp((r - \sigma^2/2)t + \sigma W_t).$$

On voit alors clairement que

$$Y_t = \frac{\partial}{\partial x}X_t(x) = \exp((r - \sigma^2/2)t + \sigma W_t) = \frac{X_t(x)}{x}.$$

Exemple 4.2. On considère le processus d'Orstein-Ulhenbeck, i.e. $b(x) = -cx$ et $\sigma(x) = \sigma$. Alors on sait que

$$X_t = xe^{-ct} + \int_0^t e^{-c(t-s)} dW_s.$$

On voit alors clairement que

$$Y_t = \frac{\partial}{\partial x}X_t(x) = e^{-ct} \quad dY_t = -cY_t dt.$$

Remarque 4.4. On peut également définir le processus dérivé, $\frac{\partial \bar{X}_t^n(x)}{\partial x}$, du schéma d'Euler $\bar{X}_t^n(x)$ de $X_t(x)$. Il est obtenu par récurrence

$$\begin{aligned} \frac{\partial \bar{X}_0^n(x)}{\partial x} &= 1 \\ \frac{\partial \bar{X}_{t_i}^n(x)}{\partial x} &= \frac{\partial \bar{X}_{t_{i-1}}^n(x)}{\partial x} + b'(\bar{X}_{t_{i-1}}^n(x)) \frac{\partial \bar{X}_{t_{i-1}}^n(x)}{\partial x} dt \\ &\quad + \sigma'(\bar{X}_{t_{i-1}}^n(x)) \frac{\partial \bar{X}_{t_{i-1}}^n(x)}{\partial x} (W_{t_i} - W_{t_{i-1}}) \end{aligned} \quad (4.3)$$

L'équation 4.1 peut être vue comme limite de 4.3 quand le pas de temps tend vers 0, $\frac{\partial \bar{X}_t^n(x)}{\partial x}$ coïncide avec le schéma d'Euler de $\frac{\partial X_t(x)}{\partial x}$.

Pour des développements plus approfondis sur ces notions, nous renvoyons à [51].

4.3.2 Applications aux calculs de couverture

Les applications se font essentiellement grâce à la proposition suivante :

Proposition 4.2. Soit g une fonction \mathcal{C}_b^1 , et X_t une diffusion vérifiant les hypothèses du théorème 4.1. Alors,

$$\frac{\partial}{\partial x} \mathbb{E}[g(X_T)] = \mathbb{E} \left[g'(X_T) \frac{\partial X_T(x)}{\partial x} \right].$$

La démonstration utilise le théorème de Lebesgue.

Muni de cet outil pour dériver sous l'espérance, on peut maintenant calculer les sensibilités. Dans la suite, nous supposons que X_T suit le modèle de Black et Scholes (i.e. $b(x) = rx$ et $\sigma(x) = \sigma x$).

L'idée générale est d'écrire les sensibilités sous la forme $\mathbb{E}[e^{-rT} g(X_T) Z_T]$ où Z_T est une variable aléatoire que l'on déterminera.

Proposition 4.3. *Lorsque X_t suit le modèle de Black Scholes, pour toute fonction $g \in \mathcal{C}_b^1$ on a*

$$\begin{aligned}\Delta &= \mathbb{E} \left[e^{-rT} g(X_T) \frac{W_T}{x\sigma T} \right] \\ \Delta &= \mathbb{E} \left[e^{-rT} g'(X_T) \frac{X_T}{x} \right]\end{aligned}$$

La première équation s'obtient sans hypothèses sur g .

Démonstration. La deuxième équation s'obtient immédiatement à partir de la proposition 4.2 et du fait que dans le modèle de Black Scholes, $X_t/x = Y_t$.

La première équation va utiliser un calcul direct. Pour éviter de dériver la fonction nous allons utiliser une intégration par parties contre la loi du Brownien.

On note $p_x(z)$ la densité de la loi de X_T . Cette fonction est régulière par rapport à x . De plus, on a

$$\mathbb{E} \left(e^{-rT} g(X_T(x)) \right) = e^{-rT} \int g(z) p_x(z) dz,$$

On dérive maintenant sous l'intégrale (théorème de Lebesgue)

$$H = e^{-rT} \int g(z) \frac{\partial p_x(z)}{\partial x} dz = \mathbb{E} \left(e^{-rT} g(X_T(x)) \frac{\frac{\partial p_x(X_T(x))}{\partial x}}{p_x(X_T(x))} \right),$$

ou

$$H = \mathbb{E} \left(e^{-rT} g(X_T(x)) \frac{\partial \log(p_x)}{\partial x}(X_T(x)) \right).$$

Dans le cas de Black Scholes, la densité de la loi de X_T est donnée par

$$p_x(z) = \frac{1}{z\sqrt{2\pi\sigma^2 T}} \exp -\frac{1}{2\sigma^2 T} \left(\log \left(\frac{z}{x} \right) - \left(r - \frac{1}{2}\sigma^2 \right) T \right)^2.$$

Ainsi,

$$\frac{\partial \log(p_x)}{\partial x}(z) = \frac{1}{x\sigma^2 T} \left(\log \left(\frac{z}{x} \right) - \left(r - \frac{1}{2}\sigma^2 \right) T \right),$$

et donc,

$$\frac{\partial \log(p_x)}{\partial x}(X_T(x)) = \frac{W_T}{x\sigma T}.$$

On conclut que

$$\Delta = \mathbb{E} \left(e^{-rT} g(X_T(x)) \frac{W_T}{x\sigma T} \right).$$

□

On peut également trouver des résultats pour le Γ et le \mathcal{V} .

Proposition 4.4. *Lorsque X_t suit le modèle de Black Scholes, pour toute fonction $g \in \mathcal{C}_b^1$ on a*

$$\Gamma = \mathbb{E} \left[e^{-rT} \left(g'(X_T) \frac{W_T X_T}{x^2 \sigma T} - g(X_T) \frac{W_T}{x^2 \sigma T} \right) \right] \quad (4.4)$$

$$\Gamma = \mathbb{E} \left[e^{-rT} \frac{g(X_T)}{x^2 \sigma T} \left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma} \right) \right] \quad (4.5)$$

$$\mathcal{V} = \mathbb{E} \left[e^{-rT} g'(X_T) X_T (W_T - \sigma T) \right] \quad (4.6)$$

$$\mathcal{V} = \mathbb{E} \left[e^{-rT} g(X_T) \left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma} \right) \right] \quad (4.7)$$

Les égalités (4.5) et (4.7) sont valables sans hypothèses de régularité sur g (borélienne bornée).

La démonstration est laissée en exercice.

Exemple 4.3. *On utilise le résultat de (4.5) pour estimer*

$$\frac{\partial^2}{\partial x^2} \mathbb{E} \mathbf{1}_{\{X_T^x \in [a, b]\}}.$$

On prend comme paramètres $x = 1$, $\sigma = 0.25$, $r = 0$, $a = 0.95$ et $b = 1.05$. On donne les intervalles de confiance obtenus par l'approche par différences finies avec différentes valeurs de ϵ et par (4.5). On fait à chaque fois 50.000 simulations. La valeur exacte est d'environ -2.53 .

Méthode	Intervalle estimé
Par (4.5)	$[-2.61, -2.51]$
Diff. finies $\epsilon = 0.5$	$[-0.32, -0.31]$
Diff. finies $\epsilon = 0.1$	$[-2.47, -2.06]$
Diff. finies $\epsilon = 0.05$	$[-3.34, -1.68]$
Diff. finies $\epsilon = 0.005$	$[-20.31, 23.91]$

On observe que le résultat est extrêmement biaisé pour $\epsilon = 0.5$ et $\epsilon = 0.1$. Pour $\epsilon = 0.005$, il est trop volatil. Même pour $\epsilon = 0.05$, le résultat est bien moins précis que celui obtenu par (4.5).

4.4 Calcul de Malliavin

Les deux derniers résultats donnent une représentation précieuse du gradient, puisqu'elle fournissent des estimateurs de Monte-Carlo très naturels. Évidemment, elles sont utiles à des fins de couverture mais peuvent également être employées dans le cadre des techniques de réduction de variance présentées dans le Chapitre 3. Toutefois, ces formulations imposent la dérivabilité du payoff (au moins dans un sens faible) et surtout la connaissance de cette dérivée. Ceci n'est pas si évident. Si l'on veut calculer le delta d'un portefeuille d'options, on ne connaît pas forcément dans le détail la forme exacte de tous les payoffs (ils sont généralement fournis par un ordinateur qui agit comme une boîte noire). Dans la Proposition 4.3, on a vu que, dans le cadre particulier du modèle de Black-Scholes, on pouvait obtenir une formulation ne faisant pas intervenir le gradient de g . Pour cela, on avait utilisé une idée d'intégration par parties par rapport à la densité gaussienne. Dans cette section, on va introduire une notion

de calcul différentiel par rapport à la trajectoire du mouvement Brownien, et une formule d'intégration par parties associée. Nous renvoyons à [46], [35] et [47] pour une présentation complète du calcul de Malliavin.

Dans cette partie pour X_t une diffusion d -dimensionnelle satisfaisant

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dW_t \\ X_0 &= x \end{aligned}$$

où W est un mouvement Brownien sur \mathbb{R}^d . Si n est le nombre de discrétisation, on a $h = T/n$ et on note $\bar{X}_{t_k}^n(x)$, $k = 0, \dots, n$ son schéma d'Euler.

4.4.1 Introduction au calcul de Malliavin

Définition 4.2. *On dira qu'une variable aléatoire F de L^2 est simple s'il existe un entier k_F , une suite $0 \leq s_1^F \dots < s_{k_F}^F \leq T$ et une fonction ϕ^F de $(\mathbb{R}^d)^{k_F}$ dans \mathbb{R} , continue, C^1 par morceaux, tels que :*

$$F = \phi^F \left(W_{s_1^F}, \dots, W_{s_{k_F}^F} \right) \quad \text{avec} \quad \sum_{j=1}^{k_F} \nabla_j \phi^F(W) \mathbf{1}_{[t, T]}(s_j^F) \in L^2 \quad \text{pour tout } t \in [0, T].$$

On note \mathcal{S} l'espace de telles fonctions.

Bien qu'il soit défini sur un espace beaucoup plus gros que \mathcal{S} , voir [46] et [48], on n'introduira ici le calcul de Malliavin que pour des fonctions simples. Il y a trois raisons à cela :

1. En général, si on sait simuler parfaitement F , c'est que $F \in \mathcal{S}$.
2. C'est beaucoup plus simple, et nous pourrons mener les preuves jusqu'au bout.
3. Discrétiser des formules obtenues en travaillant sur X ou travailler directement sur le problème discrétisé associé à \bar{X}^n revient généralement au même. On ne perd donc rien en se plaçant tout de suite dans un cadre plus simple à gérer.

Définition 4.3. *Soit $F \in \mathcal{S}$, pour tout $t \in [0, T]$, on définit*

$$D_t F := \lim_{\varepsilon \rightarrow 0} \frac{\phi^F(W + \varepsilon \mathbf{1}_{[t, T]}) - \phi^F(W)}{\varepsilon} = \sum_{j=1}^{k_F} \nabla_j \phi^F(W) \mathbf{1}_{[t, T]}(s_j^F) \quad \mathbb{P}\text{-p.s.}$$

Cette définition peut être comprise comme ceci : on choque légèrement la trajectoire du Brownien W en la remplaçant par une trajectoire $W^{\varepsilon \mathbf{1}_{[t, T]}} = W + \varepsilon \mathbf{1}_{[t, T]}$, i.e. on "shifte" le Brownien de ε après t . On regarde ensuite l'impact de ce choc sur F .

On appelle DF le processus dérivée de Malliavin.

Exemple 4.4. *On fixe $s, t \in [0, T]$.*

1. s est indépendant de W et donc $D_t s = 0$.
2. Si $F = W_s$ on a

$$D_t F = \lim_{\varepsilon \rightarrow 0} \frac{W_s + \varepsilon \mathbf{1}_{[t, T]}(s) - W_s}{\varepsilon} = \mathbf{1}_{[t, T]}(s).$$

3. Dans le modèle de Black-Scholes , on a simplement

$$D_t x e^{(r-\sigma^2/2)s+\sigma W_s} = \sigma x e^{(r-\sigma^2/2)s+\sigma W_s} \mathbf{1}_{[t,T]}(s) .$$

Remarque 4.5. $D_t F$ n'est en général pas adapté, voir l'exemple 3 ci-dessus.

Les propriétés suivantes découlent immédiatement de la définition.

Proposition 4.5. Soient F et $G \in \mathcal{S}$ et $\psi \in C_p^1$. Alors,

- (i) FG et $\psi(F) \in \mathcal{S}$,
- (ii) $D\psi(F) = \psi'(F)DF$,
- (iii) $D(FG) = (DF)G + F(DG)$.

Remarque 4.6. Si b et $\sigma \in C_b^1$, alors $\bar{X}_{t_k}^{x,n}$ est une fonction déterministe de $(W_{t_j}, 0 \leq j \leq k)$, et $\bar{X}_{t_k}^{x,n} \in \mathcal{S}$. Sa dérivée de Malliavin en t peut être calculée récursivement. C'est une matrice de taille $d \times d$.

$$\begin{aligned} (D_t \bar{X}_{t_k}^n(x))^{i,i'} &= 0 \quad \text{si } t_k < t \\ (D_t \bar{X}_{t_k}^n(x))^{i,i'} &= \sigma^{i,i'}(\bar{X}_{t_{k-1}}^n(x)) \quad \text{si } t_{k-1} < t \leq t_k \\ (D_t \bar{X}_{t_k}^n(x))^{i,i'} &= (D_t \bar{X}_{t_{k-1}}^n(x))^{i,i'} + \sum_{i''} \frac{\partial b^i}{\partial x^{i''}}(\bar{X}_{t_{i-1}}^n(x))(D_t \bar{X}_{t_{k-1}}^n(x))^{i'',i'} h \\ &\quad + \sum_{j,i''} \frac{\partial \sigma^{i,i'}}{\partial x^{i''}}(\bar{X}_{t_{k-1}}^n(x))(D_t \bar{X}_{t_{k-1}}^n(x))^{i'',i'} (W_{t_k}^j - W_{t_{k-1}}^j) \quad \text{si } t < t_{k-1} . \end{aligned} \tag{4.8}$$

Plus simplement, en dimension 1 :

$$\begin{aligned} (D_t \bar{X}_{t_k}^n(x)) &= 0 \quad \text{si } t_k < t \\ (D_t \bar{X}_{t_k}^n(x)) &= \sigma(\bar{X}_{t_{k-1}}^n(x)) \quad \text{si } t_{k-1} < t \leq t_k \\ (D_t \bar{X}_{t_k}^n(x)) &= (D_t \bar{X}_{t_{k-1}}^n(x)) + b'(\bar{X}_{t_{i-1}}^n(x))(D_t \bar{X}_{t_{k-1}}^n(x))h \\ &\quad + \sigma'(\bar{X}_{t_{k-1}}^n(x))(D_t \bar{X}_{t_{k-1}}^n(x))(W_{t_k} - W_{t_{k-1}}) \quad \text{si } t < t_{k-1} . \end{aligned}$$

Remarque 4.7. On peut observer que le gradient et la dérivée de Malliavin en t de \bar{X}^n suivent la même équation, voir remarque 4.4, à la condition en $\phi_t^n = \max \{ t_i, i = 0 \dots n \}$, tel que $t_i < t$ près

$$D_t \bar{X}_{\phi_t^n}^{x,n} = 0 \neq \nabla_x \bar{X}_{\phi_t^n}^{x,n} .$$

Si $\sigma(x)$ est inversible pour tout $x \in \mathbb{R}^d$, un calcul direct montre que

$$\nabla_x \bar{X}_{t_i}^{x,n} = D_t \bar{X}_{t_i}^{x,n} \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_x \bar{X}_{\phi_t^n}^{x,n} \quad \text{pour tout } t \leq t_i \leq T .$$

où $\phi_t^{n,+} = \max\{t_i, i = 0 \dots n, \text{ tel que } t_i \leq t\}$

L'intérêt de cette notion réside dans le fait que la stratégie de couverture d'une option peut s'écrire en fonction de la dérivée de Malliavin du payoff.

Théorème 4.6. (Formule de Clark-Ocone) Soit $F \in \mathcal{S}$, alors

$$F = \mathbb{E}F + \int_0^T \mathbb{E} [D_t F \mid \mathcal{F}_t] dW_t .$$

Démonstration. Pour simplifier, on se place dans le cas où $d = 1$. On suppose également que $\phi^F \in C_b^2$. Le cas général est obtenu par densité. A $s_{i-1}^F < t \leq s_i^F$ fixé, $\mathbb{E}[F \mid \mathcal{F}_t]$ est une fonction $\psi(t, W_t, (W_{s_j^F})_{j \leq i-1})$. On note ψ' sa dérivée par rapport au deuxième argument. On a alors pour $s_{i-1}^F < t \leq s_i^F$

$$\psi'(t, w, z) = \lim_{\varepsilon \rightarrow 0} \frac{\mathbb{E} [\phi^F(z, \tilde{\omega}^i + w + \varepsilon, \dots, \tilde{\omega}^n + w + \varepsilon) - \phi^F(z, \tilde{\omega}^i + w, \dots, \tilde{\omega}^n + w)]}{\varepsilon}$$

où $\tilde{\omega}$ est une variable aléatoire à valeurs dans \mathbb{R}^{n-i+1} distribuée selon une loi normale $\mathcal{N}(0, \Sigma)$ avec

$$\Sigma^{lk} = \min\{s_l^F - t, s_k^F - t\} .$$

On a donc par convergence dominée

$$\psi'(t, w, z) = \mathbb{E} \left[\sum_{j=i}^{k_F} \nabla_j \phi^F(z, \omega^i + w, \dots, \omega^n + w) \right]$$

d'où pour $t \in (s_{i-1}^F, s_i^F]$

$$\psi'(t, W_t, (W_{s_j^F})_{j \leq i-1}) = \mathbb{E} \left[\sum_{j=i}^{k_F} \nabla_j \phi^F(W) \mid \mathcal{F}_t \right] = \mathbb{E} [D_t \phi^F(W) \mid \mathcal{F}_t] .$$

Comme $\phi^F \in C_b^2$, on peut vérifier par des arguments similaires que ψ est $C_b^{1,2}$ par rapport à ces deux premiers arguments. En appliquant le Lemme d'Itô à la martingale $\psi(t, W_t, (W_{s_j^F})_{j \leq i-1})$ sur $(s_{i-1}^F, s_i^F]$, on obtient alors

$$\begin{aligned} \mathbb{E} [F \mid \mathcal{F}_{s_i^F}] &= \mathbb{E} [F \mid \mathcal{F}_{s_{i-1}^F}] + \int_{s_{i-1}^F}^{s_i^F} \psi'(t, W_t, (W_{s_j^F})_{j \leq i-1}) dW_t \\ &= \mathbb{E} [F \mid \mathcal{F}_{s_{i-1}^F}] + \int_{s_{i-1}^F}^{s_i^F} \mathbb{E} [D_t \phi^F(W) \mid \mathcal{F}_t] dW_t . \end{aligned}$$

Cette relation étant vérifiée pour tout $i \in \{1, \dots, k\}$, en sommant le système d'équations obtenues, on en déduit que

$$\begin{aligned} F = \mathbb{E} [F \mid \mathcal{F}_{s_k^F}] &= \mathbb{E} [F \mid \mathcal{F}_0] + \int_0^{s_k^F} \mathbb{E} [D_t \phi^F(W) \mid \mathcal{F}_t] dW_t . \\ &= \mathbb{E} [F] + \int_0^T \mathbb{E} [D_t \phi^F(W) \mid \mathcal{F}_t] dW_t . \end{aligned}$$

Théorème 4.7. (Formule d'intégration par parties) Soit $F \in \mathcal{S}$ et soit un processus adapté $h \in L^2([0, T] \times \Omega, dt \otimes dP)$, alors

$$\mathbb{E} \left[F \int_0^T h_t^* dW_t \right] = \mathbb{E} \left[\int_0^T D_t F h_t dt \right].$$

Démonstration. On pose $X_t = \mathbb{E}[F | \mathcal{F}_t]$ et $H_t = \int_0^t h_s^* dW_s$. On remarque que $H_0 = 0$. D'après le Théorème 4.6, on a donc par le Lemme d'Itô :

$$F \int_0^T h_t^* dW_t = X_T H_T = \int_0^T (X_t h_t^* + H_t \mathbb{E}[D_t F | \mathcal{F}_t]) dW_t + \int_0^T \mathbb{E}[D_t F | \mathcal{F}_t] h_t dt.$$

On obtient donc

$$\mathbb{E} \left[F \int_0^T h_t^* dW_t \right] = \mathbb{E} \left[\int_0^T \mathbb{E}[D_t F | \mathcal{F}_t] h_t dt \right] = \mathbb{E} \left[\int_0^T D_t F h_t dt \right]$$

où l'on a utilisé le Lemme de Fubini et le fait que h est adapté.

Remarque 4.8. Le Théorème 4.7 reste vrai, dans une certaine mesure, même si h n'est pas adapté. Dans ce cas, l'intégrale $\int_0^T h_t^* dW_t$ est définie en tant qu'intégrale de Skorohod, voir [46]. Lorsque $h = Fu$ où F est une variable aléatoire \mathcal{F}_T -mesurable et u est un processus adapté, on obtient, sous certaines hypothèses de régularité et d'intégrabilité, un lien entre l'intégrale d'Itô et de Skorohod :

$$\int_0^T F u_t^* dW_t = F \int_0^T u_t^* dW_t - \int_0^T D_t F u_t dt.$$

Ces résultats s'étendent au diffusion, en particulier le lien entre dérivée de Malliavin et processus tangent.

Proposition 4.8. Soit $(X_t, t \geq 0)$ une diffusion en dimension 1 dont la dynamique est fixée par l'équation différentielle stochastique

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dW_t \\ X_0 &= x \end{aligned}$$

où W est un Brownien, $b : \mathbb{R} \mapsto \mathbb{R}$ et $\sigma : \mathbb{R} \mapsto \mathbb{R}$ sont des fonctions dérivables avec des dérivées bornées. Soit Y_t le processus tangent de X_t . Y_t est solution d'une équation différentielle stochastique (voir proposition 4.1). Alors, le processus X_t est dérivable au sens de Malliavin et sa dérivée de Malliavin est donnée par

$$D_s X_t = Y_t \times Y_s^{-1} \sigma(X_s) \mathbf{1}_{s \leq t}, \quad s \geq 0 \text{ p.s.} \quad (4.9)$$

De plus, si ψ est une fonction \mathcal{C}_b^1 ,

$$D_s \psi(X_T) = \nabla \psi(X_T) Y_T \times Y_s^{-1} \sigma(X_s) \mathbf{1}_{s \leq T}, \quad s \geq 0 \text{ p.s.},$$

et

$$D_s \int_0^T \psi(X_t) dt = \int_s^T \nabla \psi(X_t) Y_t \times Y_s^{-1} \sigma(X_s) ds, \quad \text{p.s.},$$

Par ailleurs, on peut réécrire les théorèmes 4.6 et 4.7 pour les diffusions.

4.4.2 Applications aux calculs des sensibilités

Delta pour les variables simples

On commence par utiliser les résultats de la section précédente pour donner une représentation probabiliste du delta. Soit $F \in \mathcal{S}$. Alors, d'après le Théorème 4.6

$$F = \mathbb{E}[F] + \int_0^T \mathbb{E}[D_t F | \mathcal{F}_t] dW_t .$$

On fixe maintenant t et i tels que $s_{i-1}^F < t \leq s_i^F$, on a alors

$$D_t F = \sum_{j=1}^{k_F} \nabla_j \phi^F(W) \mathbf{1}_{[t, T]}(s_j^F) = \sum_{j=i}^{k_F} \nabla_j \phi^F(W)$$

qui est indépendant de $t \in (s_{i-1}^F, s_i^F]$. On a donc

$$D_t F = \int_t^{s_i^F} \frac{D_s F}{s_i^F - t} ds$$

d'où, en utilisant le Théorème 4.7

$$\mathbb{E}[D_t F | \mathcal{F}_t] = \mathbb{E} \int_t^{s_i^F} \frac{D_s F}{s_i^F - t} ds | \mathcal{F}_t = \mathbb{E} \left[F \frac{W_{s_i^F} - W_t}{s_i^F - t} | \mathcal{F}_t \right] .$$

On obtient finalement le résultat suivant

Proposition 4.9. *Soit $F \in \mathcal{S}$, alors*

$$F = \mathbb{E}[F] + \sum_{i=1}^{k_F} \int_{s_{i-1}^F}^{s_i^F} \mathbb{E} \left[F \frac{W_{s_i^F} - W_t}{s_i^F - t} | \mathcal{F}_t \right] dW_t .$$

A $s_{i-1}^F < t \leq s_i^F$ fixé, la stratégie de couverture est donc donnée par la quantité

$$\mathbb{E} \left[F \frac{W_{s_i^F} - W_t}{s_i^F - t} | \mathcal{F}_t \right]$$

, après renormalisation, qui peut être approchée par Monte-Carlo.

Delta et Gamma pour les fonctions du schéma d'Euler

On suppose maintenant que $F = g(\bar{X}_T^{x,n})$. On suppose également que $g \in C_p^1$, que σ est inversible sur \mathbb{R}^d et on pose

$$u(0, x) := \mathbb{E} [g(\bar{X}_T^{x,n})] .$$

D'après la Proposition 4.2, on a

$$\nabla u(0, x) = \mathbb{E} [\nabla g(\bar{X}_T^{x,n}) \nabla_x \bar{X}_T^{x,n}] ,$$

ce qui peut se réécrire, en utilisant la remarque 4.7 et la propriété 4.5

$$\begin{aligned}\nabla u(0, x) &= \frac{1}{T} \mathbb{E} \left[\int_0^T \nabla g(\bar{X}_T^{x,n}) D_t \bar{X}_T^{x,n} \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_x \bar{X}_{\phi_t^n}^{x,n} dt \right] \\ &= \frac{1}{T} \mathbb{E} \left[\int_0^T D_t g(\bar{X}_T^{x,n}) \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_x \bar{X}_{\phi_t^n}^{x,n} dt \right].\end{aligned}$$

En utilisant le Théorème 4.7, on obtient

$$\nabla u(0, x) = \frac{1}{T} \mathbb{E} \left[g(X_T^{x,n}) \left(\int_0^T \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_x \bar{X}_{\phi_t^n}^{x,n} \right)^* dW_t \right)^* \right].$$

En passant par densité, on obtient

Proposition 4.10. *Si $\sigma(x)$ est inversible pour tout $x \in \mathbb{R}^d$, σ^{-1} et g sont à croissance polynômiale, σ et b sont C_b^1 , alors*

$$\nabla \mathbb{E} [g(\bar{X}_T^{x,n})] = \frac{1}{T} \mathbb{E} \left[g(X_T^{x,n}) \left(\int_0^T \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_x \bar{X}_{\phi_t^n}^{x,n} \right)^* dW_t \right)^* \right].$$

On retrouve en particulier le résultat de la Proposition 4.3.

Exercice 1. *Montrer que sous les conditions de la Proposition 4.10*

$$\nabla \mathbb{E} g(X_{t_1}^{x,n}, \dots, X_{t_n}^{x,n}) = \mathbb{E} \left[g(X_{t_1}^{x,n}, \dots, X_{t_n}^{x,n}) \left(\int_0^T h_t \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_x \bar{X}_{\phi_t^n}^{x,n} \right)^* dW_t \right)^* \right],$$

où $h \in L^2([0, T], dt)$ vérifie

$$\int_0^{t_i} h_t dt = 1 \quad \text{pour tout } i \in \{1, \dots, n\}.$$

On s'intéresse maintenant au Gamma de $\mathbb{E} [g(\bar{X}_T^{x,n})]$. Si on suppose g , σ , σ^{-1} et b suffisamment régulières, on obtient en argumentant comme dans la Proposition 4.2 et en utilisant la Proposition 4.10 que :

$$\begin{aligned}\frac{\partial^2 u}{\partial x^i \partial x^j}(0, x) &= \frac{1}{T} \frac{\partial}{\partial x^j} \mathbb{E} \left[g(X_T^{x,n}) \int_0^T \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right)^* dW_t \right] \\ &= \frac{1}{T} \mathbb{E} \left[\nabla g(X_T^{x,n}) \nabla_{x^j} \bar{X}_T^{x,n} \int_0^T \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right)^* dW_t \right] \\ &+ \frac{1}{T} \mathbb{E} \left[g(X_T^{x,n}) \int_0^T \left(\sum_{k=1}^d \nabla \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \right)^k \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} \left(\nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right)^k \right)^* dW_t \right] \\ &+ \frac{1}{T} \mathbb{E} \left[g(X_T^{x,n}) \int_0^T \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right)^* dW_t \right],\end{aligned}$$

où ∇_{x^ℓ} est la dérivée par rapport à x^ℓ , i.e. $\nabla_{x^\ell} \bar{X}_{\phi_t^n}^{x,n} = \left(\nabla_x \bar{X}_{\phi_t^n}^{x,n} \right)^{\cdot \ell}$. On s'intéresse uniquement au premier terme qui fait intervenir ∇g . Pour un processus adapté h de carré intégrable, on note maintenant

$$\delta(h_t) := \int_0^T h_t^* dW_t.$$

En utilisant la Remarque 4.7 et la Propriété 4.5, on obtient

$$\begin{aligned} A &:= \mathbb{E} \left[\nabla g(X_T^{x,n}) \nabla_{x^j} \bar{X}_T^{x,n} \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \right] \\ &= \frac{1}{T} \mathbb{E} \int_0^T \nabla g(\bar{X}_T^{x,n}) D_t \bar{X}_T^{x,n} \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} dt \\ &= \frac{1}{T} \mathbb{E} \left[\int_0^T D_t \left(g(\bar{X}_T^{x,n}) \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \right) \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} dt \right] \\ &\quad - \frac{1}{T} \mathbb{E} \left[\int_0^T g(\bar{X}_T^{x,n}) D_t \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} dt \right]. \end{aligned}$$

D'après le Théorème 4.7, ceci implique que

$$\begin{aligned} A &= \frac{1}{T} \mathbb{E} \left[g(\bar{X}_T^{x,n}) \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} \right) \right] \\ &\quad - \frac{1}{T} \mathbb{E} \left[\int_0^T g(\bar{X}_T^{x,n}) D_t \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} dt \right]. \end{aligned}$$

En regroupant les termes et en passant par densité, on obtient

Proposition 4.11. *Si σ est inversible sur \mathbb{R}^d , $a^{-1} \in C_p^1$, g est à croissance polynômiale, σ et b sont C_b^2 , alors*

$$\frac{\partial}{\partial x^i \partial x^j} \mathbb{E} [g(\bar{X}_T^{x,n})] = \mathbb{E} [g(X_T^{x,n}) \Pi^{ij}],$$

où

$$\begin{aligned} \Pi^{ij} &:= \frac{1}{T^2} \left(\delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} \right) \right) \\ &\quad - \frac{1}{T^2} \left(\int_0^T D_t \delta \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right) \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} dt \right) \\ &\quad + \frac{1}{T} \delta \left(\sum_{k=1}^d \nabla \left(\sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \right)^{\cdot k} \nabla_{x^j} \bar{X}_{\phi_t^n}^{x,n} \left(\nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right)^{\cdot k} + \sigma(\bar{X}_{\phi_t^n}^{x,n})^{-1} \nabla_{x^j} \nabla_{x^i} \bar{X}_{\phi_t^n}^{x,n} \right). \end{aligned}$$

Exercice 2. Vérifier que dans le modèle de Black-Scholes $\nabla_{x^i} \nabla_{x^j} X^x = 0$ et retrouver ainsi la formule (4.4).

Exercice 3. Étendre la formule de la Proposition 4.11 à des payoffs de la forme

$$g(\bar{X}_{t_1}^{x,n}, \dots, \bar{X}_{t_n}^{x,n})$$

sur le modèle de l'Exercice 1.

Remarque 4.9. Cette approche a été initiée par [24]. Les formules que nous démontrons ici correspondent à la discrétisation des formules obtenues en considérant un payoff dépendant de X au lieu de \bar{X}^n . L'étude de ces méthodes a été poursuivie dans [25], voir également [18]. Des extensions au cas des options barrière et lookback ont été obtenues par [29]. On pourra également consulter [10, 9, 11] qui traite notamment le cas des options asiatiques et [34] qui développe des approches alternatives.

Chapitre 5

Les options américaines

Le calcul du prix et de la couverture d'une option américaine est un sujet d'actualité en finance. Le but est d'essayer de s'affranchir du problème de la dimension. En effet, l'un des points forts des méthodes de Monte Carlo est d'être moins dépendant de la dimension du problème que les méthodes liées aux équations aux dérivées partielles.

Nous présentons des développements récents d'algorithmes de Monte Carlo pour calculer le prix d'une option américaine. Il existe trois classes d'algorithmes :

1. La première utilise des approximations de l'espérance conditionnelle (Longstaff et Schwartz [43] et Tsitsiklis et VanRoy [59]) : nous traiterons uniquement Longstaff Schwartz.
2. La deuxième est fondée sur l'approximation du processus de Markov sous-jacent (algorithme de quantification [5, 49] (voir le cours de H. Pham), Broadie et Glassermann [14, 15] et Barraquand et Martineau [8]) que l'on n'abordera pas ici.
3. La troisième utilise le Calcul de Malliavin pour calculer numériquement l'espérance conditionnelle [25, 12, 13, 20].

5.1 Introduction

Dans ce chapitre, nous étudierons les options Bermudéennes qui sont des options américaines ne pouvant être exercées qu'à certaines dates fixes $(t_0, \dots, t_N = T)$.

Le sous-jacent au temps $(t_n, 0 \leq n \leq N)$ est une chaîne de Markov avec comme matrice de transition de n à $n + 1$ $P_n(x, dy)$. Cela signifie que pour toute fonction f

$$\mathbb{E}(f(X_{n+1})|\mathcal{F}_n) = P_n f(X_n) := \int_{\mathbb{R}} f(y)P_n(X_n, dy),$$

où $\mathcal{F}_n = \sigma(X_k, k \leq n)$.

On suppose que le payoff au temps t_n est donné par $\varphi(X_n)$ où φ est une fonction bornée et que le facteur d'actualisation entre t_j et t_{j+1} s'exprime par $1/(1+r_j)$, r_j étant une constante strictement positive représentant le taux d'intérêt entre les temps t_j et t_{j+1} . Cela permet de définir un facteur d'actualisation entre les temps j et k , quand $j < k$, en posant

$$B(j, k) = \prod_{l=j}^{k-1} \frac{1}{1+r_l}.$$

On remarque que $B(j, j+1) = 1/(1+r_j)$. Utilisant ces notations, le prix à $t=0$ de cette option est

$$Q_0 = \sup_{\tau \in \mathcal{T}_{0,N}} \mathbb{E}[B(0, \tau) \varphi(X_\tau)] \quad (5.1)$$

où $\mathcal{T}_{0,N}$ est l'ensemble des \mathcal{F}_n -temps d'arrêt prenant leurs valeurs dans $\{0, \dots, N\}$.

Remarque 5.1. *Le modèle de Black Scholes à d dimensions $S_t^{x,i}, 1 \leq i \leq d$, peut s'écrire*

$$S_t^{x,i} = x_i e^{(r - \frac{1}{2}\sigma_i^2)t + \sum_{1 \leq j \leq p} \sigma_{ij} W_t^j}$$

où $(W_t, t \geq 0)$ est un mouvement Brownien r -dimensionnel, $\sigma_i^2 = \sum_{1 \leq j \leq p} \sigma_{ij}^2$ et $r > 0$ le taux d'intérêt sans risque. La matrice de transition est définie par

$$P_n f(x) = \mathbb{E}\left(f(S_{t_{n+1}}^x) | S_{t_n} = x\right) = \mathbb{E}\left(f(S_{t_{n+1}-t_n}^x)\right).$$

De plus, on a

$$1 + r_n = \exp(r(t_{n+1} - t_n)).$$

Remarque 5.2. *Le schéma d'Euler d'une diffusion d -dimensionnelle peut s'écrire*

$$\bar{X}_{t_{k+1}}^n = \bar{X}_{t_k}^n + b(\bar{X}_{t_k}^n)h + \sigma(\bar{X}_{t_k}^n) \cdot (W_{t_{k+1}} - W_{t_k})$$

où $(W_t, t \geq 0)$ est un mouvement Brownien r -dimensionnel et $r > 0$ le taux d'intérêt sans risque. La matrice de transition est définie par

$$P_k f(x) = \mathbb{E}\left(f(\bar{X}_{t_{k+1}}^n) | \bar{X}_{t_k}^n = x\right) = \mathbb{E}\left(f(x + b(x)h + \sigma(x) \cdot (W_{t_{k+1}} - W_{t_k}))\right).$$

De plus, on a

$$1 + r_k = \exp(rh).$$

Attention, nous avons montré que la suite X_{t_1}, \dots, X_{t_n} est une chaîne de Markov. Mais le schéma d'Euler n'est pas un processus de Markov.

Les théories de contrôle stochastique montre que $Q_0 = u(0, X_0)$, où u est la solution de l'algorithme de programmation dynamique

$$\begin{cases} u(N, x) = \varphi(x) \\ u(j, x) = \max_{x \in \mathbb{R}^d} (\varphi(x), B(j, j+1)P_j u(j+1, x)), \quad 1 \leq j \leq N. \end{cases} \quad (5.2)$$

Le temps d'arrêt

$$\tau^* = \inf \{j \geq 0, u(j, X_j) = \varphi(X_j)\}, \quad (5.3)$$

est optimal, i.e.

$$Q_0 = u(0, X_0) = \mathbb{E}(B(0, \tau^*) \varphi(X_{\tau^*})).$$

De plus, le prix Q_j de l'option américaine au temps j donné par

$$Q_j = \sup_{\tau \in \mathcal{T}_{j,N}} \mathbb{E}(B(j, \tau) \varphi(X_\tau) | \mathcal{F}_j),$$

peut être calculé comme $u(j, X_j)$ et le temps d'arrêt optimal au temps j sera donné par τ_j^* où

$$\tau_j^* = \inf \{i \geq j; u(i, X_i) = \varphi(X_i)\}.$$

5.2 L'algorithme de Longstaff Schwartz

L'algorithme de Longstaff Schwartz consiste en une approximation du temps d'arrêt optimal τ^* sur M trajectoires, par les variables aléatoires $\tau^{(m)}$, qui dépendent de la trajectoire m . Le prix est ensuite approché en utilisant une méthode de Monte Carlo :

$$Q_0^M = \frac{1}{M} \sum_{1 \leq i \leq M} B(0, \tau^{(m)}) \varphi \left(X_{\tau^{(m)}}^{(m)} \right).$$

5.2.1 Une équation de programmation dynamique pour τ^*

La principale caractéristique de cet algorithme est d'utiliser un principe de programmation pour le temps d'arrêt optimal lui-même et non pour la fonction valeur.

Si on remarque que $\tau^* = \tau_0^*$, le temps d'arrêt optimal au temps 0, peut être calculé en utilisant la suite $(\tau_j^*, 0 \leq j \leq N)$ définie par la récurrence suivante :

$$\begin{cases} \tau_N^* = N, \\ \tau_j^* = j \mathbf{1}_{\{\varphi(X_j) \geq u(j, X_j)\}} + \tau_{j+1}^* \mathbf{1}_{\{\varphi(X_j) < u(j, X_j)\}}. \end{cases} \quad (5.4)$$

De plus, on a

$$\begin{aligned} \{\varphi(X_j) < u(j, X_j)\} &= \{\varphi(X_j) < B(j, \tau_{j+1}^*) P_j u(j+1, X_j)\} \\ &= \left\{ \varphi(X_j) < \mathbb{E} \left(B(j, \tau_{j+1}^*) \varphi(X_{\tau_{j+1}^*}) \middle| X_j \right) \right\}. \end{aligned}$$

Avec cette définition pour τ_j^* , il est facile de vérifier récursivement que

$$\tau_j^* = \min \{i \geq j, \varphi(X_i) = u(i, X_i)\}. \quad (5.5)$$

Ainsi, les τ_j^* (et donc $\tau^* = \tau_0^*$) sont les temps d'arrêt optimaux au temps j . Afin d'estimer numériquement τ^* nous devons trouver une valeur approchée de l'espérance conditionnelle

$$\mathbb{E} \left(B(j, \tau_{j+1}^*) \varphi \left(X_{\tau_{j+1}^*} \right) \middle| X_j \right). \quad (5.6)$$

L'idée fondamentale de Longstaff Schwartz est d'introduire une méthode de moindre carré pour calculer (5.6).

5.2.2 La méthode de régression

On note Z_{j+1} la variable aléatoire

$$Z_{j+1} = B(j, \tau_{j+1}^*) \varphi \left(X_{\tau_{j+1}^*} \right).$$

Comme B et φ sont bornées, Z_{j+1} l'est aussi. L'espérance conditionnelle peut aussi être vue comme une projection dans L^2 , donc

$$\mathbb{E}(Z_{j+1} | X_j)$$

peut s'exprimer comme $\psi_j(X_j)$, où ψ_j minimise

$$\mathbb{E} \left([X_{j+1} - f(X_j)]^2 \right)$$

parmi les fonctions f telle que $\mathbb{E} (f(X_j)^2) < +\infty$.

On se donne maintenant une suite de fonctions $(g_l, l \geq 1)$ qui est une base de $L^2 = L^2(\mathbb{R}^d, \text{loi de } X_j)$ pour tous $j, 1 \leq j \leq N$. Pour tous j , ψ_j peut s'exprimer comme

$$\psi_j(x) = \sum_{l \geq 1} \alpha_l g_l(x),$$

où la convergence de la série s'effectue dans L^2 . Nous avons donc une manière de calculer le temps d'arrêt optimal en utilisant la procédure :

1. On initialise $\tau_N = N$.
2. On définit $\alpha^j = (\alpha_l^j, l \geq 0)$ comme la suite minimisant

$$\mathbb{E} \left(\left[B(j, \tau_{j+1}^*) \varphi(X_{\tau_{j+1}^*}) - (\alpha \cdot g)(X_j) \right]^2 \right)$$

où $\alpha \cdot g = \sum_{l \geq 1} \alpha_l g_l$.

3. On définit $\tau_j^* = j \mathbf{1}_{\varphi(X_j) \geq (\alpha^j \cdot g)(X_j)} + \tau_{j+1}^* \mathbf{1}_{\varphi(X_j) < (\alpha^j \cdot g)(X_j)}$.

Cet algorithme n'est pas réellement utilisable car il est impossible de calculer l'espérance du problème de minimisation au pas 2.

Pour implémenter cet algorithme, l'idée est de tronquer la série à l'indice k . Cela conduit à une procédure modifiée :

1. On initialise $\hat{\tau}_N = N$. Alors,
2. On définit $\hat{\alpha}^{j,k} = (\hat{\alpha}_l^{j,k}, 0 \leq l \leq k)$ comme le vecteur minimisant

$$\mathbb{E} \left(\left[B(j, \hat{\tau}_{j+1}) \varphi(X_{\hat{\tau}_{j+1}}) - (\hat{\alpha}^{j,k} \cdot g)(X_j) \right]^2 \right) \quad (5.7)$$

où $(\hat{\alpha}^{j,k} \cdot g)(x) = \sum_{l=1}^k \hat{\alpha}_l^{j,k} g_l(x)$.

3. On définit

$$\hat{\tau}_j = j \mathbf{1}_{\{\varphi(X_j) \geq (\hat{\alpha}^{j,k} \cdot g)(X_j)\}} + \hat{\tau}_{j+1} \mathbf{1}_{\{\varphi(X_j) < (\hat{\alpha}^{j,k} \cdot g)(X_j)\}}.$$

Enfin, il reste à remplacer l'espérance de (5.7) par du Monte Carlo.

5.2.3 Du Monte Carlo pour le problème de régression

Soit $(X_n^{(m)}, 0 \leq n \leq N)$, pour $1 \leq m \leq M$, M trajectoires simulant le processus $(X_n, 0 \leq n \leq N)$. On remplace le problème de minimisation 5.7 par

1. On initialise $\tau_N^M = N$.
2. On définit $\alpha_M^{j,k} = (\alpha_M^{j,k}, 0 \leq j \leq k)$ comme le vecteur minimisant

$$\frac{1}{M} \sum_{1 \leq m \leq M} \left((\alpha \cdot g)(X_j^{(m)}) - B(j, \tau_{j+1}) \varphi(X_{\tau_{j+1}}^{(m)}) \right)^2 \quad (5.8)$$

3. On définit pour chaque trajectoire m

$$\tau_j^{(m)} = j \mathbf{1}_{\varphi(X_j) \geq (\alpha_M^{j,k} \cdot g)(X_j)} + \tau_{j+1}^{(m)} \mathbf{1}_{\varphi(X_j) < (\alpha_M^{j,k} \cdot g)(X_j)}.$$

Cet algorithme est maintenant fonctionnel et l'estimateur du prix est donné par

$$\frac{1}{M} B(0, \tau^{(m)}) \varphi(X_{\tau^{(m)}}).$$

Remarque 5.3. – *Le problème de minimisation (5.8) est un problème standard d'estimation par les moindres carrés. Voir par exemple dans [50].*

- *La base $(g_k, k \geq 1)$ est supposée indépendante du temps j . Cela est juste par simplicité et on peut très bien changer de base à chaque temps.*
- *Très souvent, en finance, pour le call et le put*

$$\mathbb{P}(\varphi(X_j) = 0) > 0,$$

pour tous j . Comme $\{\varphi(X_j) = 0\} \subset A_j$, il n'est pas nécessaire de calculer

$$\mathbb{E} \left(B(j, \tau_{j+1}^*) \varphi(X_{\tau_{j+1}^*}) \middle| X_j \right)$$

sur l'ensemble $\{\varphi(X_j) = 0\}$ et on peut se restreindre aux trajectoires telles que $\left\{ \varphi(X_j) > 0 \right\}$.

- *Une preuve rigoureuse de la convergence de cet algorithme est donnée dans [17].*

5.3 Approximation d'espérances conditionnelles

L'approximation d'espérances conditionnelles à l'aide du calcul de Malliavin a été initiée par [25], puis par [12, 13, 20]. Nous allons ici essayer de donner un aperçu de ces méthodes. Par souci de simplifier l'écriture nous nous limiterons au cas $d = 1$. Évidemment, l'intérêt de ces méthodes est justement le cas $d > 1$ car en dimension 1, des méthodes déterministes fournissent de bon moyen pour calculer le prix d'une option américaine. Pour plus de détails sur le cas général, on pourra consulter [4].

On considère donc une diffusion réelle satisfaisant l'équation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \quad X_0 = x \in \mathbb{R}.$$

On suppose que b et σ sont dans \mathcal{C}_b^1 . On note $\Delta^n W_i = W_{t_{i+1}} - W_{t_i}$. On lui associe son schéma d'Euler $\bar{X}_t^n(x)$. Le but est donc de calculer

$$\mathbb{E} \left[f(\bar{X}_{t_{k+1}}^n) \middle| \bar{X}_{t_k}^n = y \right].$$

où $y \in \mathbb{R}$ et f une fonction à croissance polynômiale. Dans ce cas on a le théorème

5.3.1 Calcul de Malliavin : compléments

Le point central de la théorie du Calcul de Malliavin est la formule d'intégration par parties.

Définition 5.1. Soient F et G deux variables aléatoires réelles de carrés intégrables. On dit qu'elles vérifient la propriété $IP(F, G)$ ("Integration by parts") si il existe une variable aléatoire de carré intégrable telle que :

$$\mathbb{E} [\phi'(F)G] = \mathbb{E} [\phi(F)\pi_F(G)], \quad \forall \phi \in C_b^\infty(\mathbb{R}, \mathbb{R}).$$

On a alors le lemme suivant

Lemme 5.1. 1. On suppose que $IP(F, 1)$ est vraie. Alors la loi de la variable aléatoire F a une densité p donnée par

$$p(\alpha) = \mathbb{E} [H(F - \alpha)\pi_F(1)],$$

où $H(x) = \mathbf{1}_{x \geq 0}$.

2. Si $IP(F, 1)$ et $IP(F, G)$ sont vraies, alors

$$\mathbb{E} [G | F = \alpha] = \frac{\mathbb{E} [H(F - \alpha)\pi_F(G)]}{\mathbb{E} [H(F - \alpha)\pi_F(1)]}.$$

Remarque 5.4. Cette formulation de la densité a été abondamment utilisée par [46] pour étudier la régularité des densités de variables aléatoires. Une étude sur les méthodes de réduction de variance de l'estimateur ainsi obtenu a été menée par [21] en dimension 1.

Démonstration. 1. Nous allons employer des techniques de régularisation. Soit une fonction régulière, symétrique, positive ϕ dont le support est contenu dans $[-1, 1]$ et telle que $\int_{-1}^1 \phi(t)dt = 1$. On considère alors $\phi_\delta(x) = \phi(x/\delta)/\delta$ et $\Phi_\delta(x) = \int_{-\infty}^x \phi_\delta(t)dt$. On remarque que $\Phi'_\delta(x) = \phi_\delta(x)$ et que $\lim_{\delta \rightarrow 0} \Phi_\delta(x) = \tilde{\delta}(x)$ où

$$\begin{aligned} \tilde{\delta}(x) &= 0 \quad \text{si } x < 0 \\ &= 1/2 \quad \text{si } x = 0 \\ &= 1 \quad \text{si } x > 0. \end{aligned}$$

Soit maintenant G_δ une variable aléatoire indépendante de F telle que sa loi est $\phi_\delta(x)dx$. Alors, $\lim_{\delta \rightarrow 0} \mathbb{E} [f(F - G_\delta)] = \mathbb{E} [f(F)]$ pour toute fonction f continue. Ainsi,

$$\mathbb{E} [f(F - G_\delta)] = \int_{\mathbb{R}} \mathbb{E} [f(F - z)\phi_\delta(z)]dz = \int_{\mathbb{R}} f(z)\mathbb{E} [\phi_\delta(F - z)]dz.$$

Comme la propriété $IP(F, 1)$ est vraie

$$\mathbb{E} [\phi_\delta(F - z)] = \mathbb{E} [\Phi'_\delta(F - z)] = \mathbb{E} [\Phi_\delta(F - z)\pi_F(1)].$$

D'où d'après le théorème de Lebesgue

$$\lim_{\delta \rightarrow 0} \mathbb{E} [f(F - G_\delta)] = \int_{\mathbb{R}} f(z)\mathbb{E} [\tilde{\delta}(F - z)\pi_F(1)] dz.$$

Donc pour toute fonction f continue,

$$\mathbb{E}[f(F)] = \int_{\mathbb{R}} f(z) \mathbb{E}[\tilde{\delta}(F-z)\pi_F(1)] dz.$$

ce qui prouve que la densité de F est donnée par

$$p(z) = \mathbb{E}[\tilde{\delta}(F-z)\pi_F(1)] = \mathbb{E}[\mathbf{1}_{[z,+\infty)}(F)\pi_F(1)] = \mathbb{E}[H(F-z)\pi_F(1)].$$

On a pu remplacer $\tilde{\delta}(x)$ par $\mathbf{1}_{[0,+\infty)}$ car on sait déjà que la loi de F est absolument continue par rapport à la mesure de Lebesgue.

2. Nous avons juste à vérifier que pour toute fonction f continue bornée on a $\mathbb{E}[f(F)G] = \mathbb{E}[f(F)\theta(G)]$ où

$$\theta(z) = \frac{\mathbb{E}[H(F-z)\pi_F(G)]}{\mathbb{E}[H(F-z)\pi_F(1)]}.$$

En utilisant les mêmes fonctions de régularisation que ci-dessus la propriété $IP(F, G)$ permet d'affirmer que

$$\begin{aligned} \mathbb{E}[f(F)G] &= \mathbb{E}\left[G \lim_{\delta \rightarrow 0} \int_{\mathbb{R}} f(z) \phi_{\delta}(F-z) dz\right] = \lim_{\delta \rightarrow 0} \int_{\mathbb{R}} f(z) \mathbb{E}[G \phi_{\delta}(F-z)] dz \\ &= \lim_{\delta \rightarrow 0} \int_{\mathbb{R}} f(z) \mathbb{E}[G \Phi'_{\delta}(F-z)] dz \\ &= \lim_{\delta \rightarrow 0} \int_{\mathbb{R}} f(z) \mathbb{E}[G \Phi_{\delta}(F-z) \pi_F(G)] dz \\ &= \int_{\mathbb{R}} f(z) \mathbb{E}[G \mathbf{1}_{[0, \infty)}(F-z) \pi_F(G)] dz \\ &= \int_{\mathbb{R}} f(z) \theta(z) p(z) dz = \mathbb{E}[f(F)\theta(F)]. \end{aligned}$$

□

Nous allons également donner une version localisée du lemme ci-dessus.

Définition 5.2. Soient F et G deux variables aléatoires réelles de carrés intégrables et soit $D \subset \mathbb{R}$. On dit qu'elles vérifient la propriété $IP_D(F, G)$ ("Integration by parts") si il existe une variable aléatoire de carré intégrable telle que :

$$\mathbb{E}[\phi'(F)G] = \mathbb{E}[\phi(F)\pi_F^D(G)], \quad \forall \phi \in \mathcal{C}_b^{\infty}(\mathbb{R}, \mathbb{R}).$$

On retrouve alors le même lemme dont la preuve est rigoureusement identique.

Lemme 5.2. 1. On suppose que $IP_D(F, 1)$ est vraie. Alors la loi de la variable aléatoire F a une densité p locale sur D donnée par

$$p_D(\alpha) = \mathbb{E}[H(F-\alpha)\pi_F^D(1)].$$

2. Si $IP_D(F, 1)$ et $IP_D(F, G)$ sont vraies, alors

$$\mathbb{E}[G \mid F = \alpha] = \frac{\mathbb{E}[H(F - \alpha)\pi_F(G)]}{\mathbb{E}[H(F - \alpha)\pi_F(1)]} \quad \forall \alpha \in D.$$

Nous devons maintenant relier les définitions 5.1 et 5.2 à la formule d'intégration par parties établies précédemment 4.7 et au schéma d'Euler.

Lemme 5.3. Soit $\bar{X}_{t_i}^n$ le schéma d'Euler d'une diffusion en dimension 1 et g une fonction $\mathbb{R} \rightarrow \mathbb{R}$ à croissance polynômiale. Alors

1. La propriété $IP(\bar{X}_{t_i}^n, g(\bar{X}_{t_{i+1}}^n))$ est vraie, i.e.

$$\mathbb{E}[\phi'(\bar{X}_{t_i}^n)g(\bar{X}_{t_{i+1}}^n)] = \mathbb{E}[\phi(\bar{X}_{t_i}^n)\pi_{\bar{X}_{t_i}^n}[g](\bar{X}_{t_{i+1}}^n)],$$

avec

$$\begin{aligned} \pi_{\bar{X}_{t_i}^n}[g](\bar{X}_{t_{i+1}}^n) &= g(\bar{X}_{t_{i+1}}^n) \left(\frac{\Delta^n W_i}{h\sigma(\bar{X}_{t_{i-1}}^n)} \right. \\ &\quad \left. - \frac{1}{h\sigma(\bar{X}_{t_i}^n)} \left((1 + b'(\bar{X}_{t_i}^n)h) \Delta^n W_{i+1} + \sigma'(\bar{X}_{t_i}^n)(\Delta^n W_{i+1}^2 - h) \right) \right). \end{aligned} \quad (5.9)$$

2. Pour un $\alpha > 0$ fixe, soit $\phi \in C_b^1(\mathbb{R})$ telle que $\text{Supp}(\phi') \subset B_\varepsilon(\alpha) = (\alpha - \varepsilon, \alpha + \varepsilon)$, avec $\varepsilon > 0$. Soit $\psi \in C_b^1(\mathbb{R})$ telle que $\psi|_{B_\varepsilon(\alpha)} = 1$. Alors,

$$\mathbb{E}[\phi'(\bar{X}_{t_i}^n)g(\bar{X}_{t_{i+1}}^n)] = \mathbb{E}[\phi(\bar{X}_{t_i}^n)\pi_{\bar{X}_{t_i}^n}^\psi[g](\bar{X}_{t_{i+1}}^n)],$$

avec

$$\begin{aligned} \pi_{\bar{X}_{t_i}^n}^\psi[g](\bar{X}_{t_{i+1}}^n) &= g(\bar{X}_{t_{i+1}}^n) \left(-\psi'(\bar{X}_{t_i}^n) + \psi(\bar{X}_{t_i}^n) \left[\frac{\Delta^n W_i}{h\sigma(\bar{X}_{t_{i-1}}^n)} \right. \right. \\ &\quad \left. \left. - \frac{1}{h\sigma(\bar{X}_{t_i}^n)} \left((1 + b'(\bar{X}_{t_i}^n)h) \Delta^n W_{i+1} + \sigma'(\bar{X}_{t_i}^n)(\Delta^n W_{i+1}^2 - h) \right) \right] \right). \end{aligned} \quad (5.10)$$

En particulier $IP_D(\bar{X}_{t_i}^n, g(\bar{X}_{t_{i+1}}^n))$ est vraie avec $D = B_\varepsilon(\alpha)$

Démonstration. Nous allons faire la preuve du point 1, celle du point 2 s'obtenant avec les mêmes arguments (remarquer que $\phi' = \phi'\psi$). La preuve va consister en l'application de la formule d'intégration par parties. Pour cela on doit d'abord remarquer que

$$\sigma(\bar{X}_{t_i}^n)h = \int_{t_i}^{t_{i+1}} \sigma(\bar{X}_{t_i}^n)dt = \int_{t_i}^{t_{i+1}} D_t \bar{X}_{t_i}^n dt. \quad (5.11)$$

Ensuite, on s'aperçoit que

$$g(\bar{X}_{t_{i+1}}^n) = g(\bar{X}_{t_i}^n + b(\bar{X}_{t_i}^n)h + \sigma(\bar{X}_{t_i}^n)\Delta^n W_{i+1}) := g(\delta(\bar{X}_{t_i}^n, \Delta^n W_{i+1})).$$

Comme $\Delta^n W_{i+1}$ est indépendant de \mathcal{F}_{t_i} ,

$$\mathbb{E} [\phi'(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, \Delta^n W_{i+1}))] = \mathbb{E} \left[\mathbb{E} [\phi'(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y))] \Big|_{y=\Delta^n W_{i+1}} \right].$$

Étape 1 : Nous allons commencer par calculer $D_t (\phi(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y)))$.

$$\begin{aligned} \phi'(\bar{X}_{t_i}^n)D_t \bar{X}_{t_i}^n g(\delta(\bar{X}_{t_i}^n, y)) &= D_t (\phi(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y))) \\ &\quad - \phi'(\bar{X}_{t_i}^n)g'(\delta(\bar{X}_{t_i}^n, y))(1 + b'(\bar{X}_{t_i}^n)h + \sigma'(\bar{X}_{t_i}^n)y)D_t \bar{X}_{t_i}^n \end{aligned}$$

Utilisant la remarque (5.11) et Fubini, on a

$$\begin{aligned} \mathbb{E} [\phi'(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y))] &= \mathbb{E} [D_t (\phi(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y)))] \\ &\quad - \mathbb{E} [\phi(\bar{X}_{t_i}^n)g'(\delta(\bar{X}_{t_i}^n, y))(1 + b'(\bar{X}_{t_i}^n)h + \sigma'(\bar{X}_{t_i}^n)y)D_t \bar{X}_{t_i}^n] \\ &= \underbrace{\frac{1}{\sigma(\bar{X}_{t_{i-1}}^n)h} \mathbb{E} \left[\int_{t_{i-1}}^{t_i} D_t (\phi(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y))) dt \right]}_A \\ &\quad - \underbrace{\mathbb{E} [\phi(\bar{X}_{t_i}^n)g'(\delta(\bar{X}_{t_i}^n, y))(1 + b'(\bar{X}_{t_i}^n)h)]}_B \\ &\quad - \underbrace{\mathbb{E} [\phi(\bar{X}_{t_i}^n)g'(\delta(\bar{X}_{t_i}^n, y))\sigma'(\bar{X}_{t_i}^n)y]D_t \bar{X}_{t_i}^n}_C \end{aligned}$$

Étape 2 : le terme A On applique la formule d'intégration par parties 4.7 conditionnée à $\mathcal{F}_{t_{i-1}}$ ce qui donne

$$\begin{aligned} A &= \frac{1}{\sigma(\bar{X}_{t_{i-1}}^n)h} \mathbb{E} \left[\phi(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y)) \int_{t_{i-1}}^{t_i} dW_t \right] \\ &= \frac{1}{\sigma(\bar{X}_{t_{i-1}}^n)h} \mathbb{E} [\phi(\bar{X}_{t_i}^n)g(\delta(\bar{X}_{t_i}^n, y))\Delta^n W_i]. \end{aligned}$$

On a donc obtenu l'identité :

$$\mathbb{E} [\phi'(\bar{X}_{t_i}^n)|\mathcal{F}_{t_{i-1}}] = \frac{1}{\sigma(\bar{X}_{t_{i-1}}^n)h} \mathbb{E} [\phi(\bar{X}_{t_i}^n)\Delta^n W_i|\mathcal{F}_{t_{i-1}}] \quad (5.12)$$

Étape 3 : le terme B En fait ce que l'on veut calculer est

$$\begin{aligned} \mathbb{E} [B|_{y=\Delta^n W_{i+1}}] &= \mathbb{E} [\phi(\bar{X}_{t_i}^n)g'(\bar{X}_{t_{i+1}}^n)(1 + b'(\bar{X}_{t_i}^n)h)] \\ &= \mathbb{E} [\phi(\bar{X}_{t_i}^n)(1 + b'(\bar{X}_{t_i}^n)h)\mathbb{E} [g'(\bar{X}_{t_{i+1}}^n)|\mathcal{F}_{t_i}]]. \end{aligned}$$

En appliquant (5.12) à g et au temps t_i on a

$$\mathbb{E} [B|_{y=\Delta^n W_{i+1}}] = \frac{1}{\sigma(\bar{X}_{t_i}^n)h} \mathbb{E} [\phi(\bar{X}_{t_i}^n)(1 + b'(\bar{X}_{t_i}^n)h)g(\bar{X}_{t_{i+1}}^n)\Delta^n W_{i+1}].$$

Étape 4 : le terme C On commence comme pour B

$$\begin{aligned}\mathbb{E} [C|_{y=\Delta^n W_{i+1}}] &= \mathbb{E} \left[\phi(\bar{X}_{t_i}^n) g'(\bar{X}_{t_{i+1}}^n) \sigma'(\bar{X}_{t_i}^n) \Delta^n W_{i+1} \right] \\ &= \mathbb{E} \left[\phi(\bar{X}_{t_i}^n) \sigma'(\bar{X}_{t_i}^n) \mathbb{E} \left[g'(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1} | \mathcal{F}_{t_i} \right] \right].\end{aligned}$$

Or $D_t \Delta^n W_{i+1} = 1$ pour $t \in (t_i, t_{i+1}]$, et

$$D_t \left(g(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1} \right) = g'(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1} D_t \bar{X}_{t_{i+1}}^n + g(\bar{X}_{t_{i+1}}^n).$$

La formule d'intégration par parties permet alors de dire que

$$\begin{aligned}\mathbb{E} \left[g'(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1} | \mathcal{F}_{t_i} \right] &= \frac{1}{\sigma(\bar{X}_{t_i}^n) h} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} D_t \bar{X}_{t_{i+1}}^n g'(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1} dt | \mathcal{F}_{t_i} \right] \\ &= \frac{1}{\sigma(\bar{X}_{t_i}^n) h} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} D_t \left(g(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1} \right) dt | \mathcal{F}_{t_i} \right] \\ &\quad - \frac{1}{\sigma(\bar{X}_{t_i}^n)} \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) | \mathcal{F}_{t_i} \right] \\ &= \frac{1}{\sigma(\bar{X}_{t_i}^n) h} \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1} \int_{t_i}^{t_{i+1}} dW_t | \mathcal{F}_{t_i} \right] \\ &\quad - \frac{1}{\sigma(\bar{X}_{t_i}^n)} \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) | \mathcal{F}_{t_i} \right] \\ &= \frac{1}{\sigma(\bar{X}_{t_i}^n) h} \left(\mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) \Delta^n W_{i+1}^2 | \mathcal{F}_{t_i} \right] - h \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) | \mathcal{F}_{t_i} \right] \right)\end{aligned}$$

Ainsi,

$$\mathbb{E} [C|_{y=\Delta^n W_{i+1}}] = \frac{1}{\sigma(\bar{X}_{t_i}^n) h} \mathbb{E} \left[\phi(\bar{X}_{t_i}^n) \sigma'(\bar{X}_{t_i}^n) g(\bar{X}_{t_{i+1}}^n) (\Delta^n W_{i+1}^2 - h) \right].$$

En rassemblant les trois termes on obtient la preuve du théorème. \square

On peut maintenant énoncer le théorème principal

Théorème 5.4. 1. Pour tous $k = 0, \dots, n-1$, et $y \in \mathbb{R}$, on a

$$\mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) | \bar{X}_{t_i}^n = y \right] = \frac{\mathbb{E} \left[H(\bar{X}_{t_i}^n - y) \pi_{\bar{X}_{t_i}^n} [g](\bar{X}_{t_{i+1}}^n) \right]}{\mathbb{E} \left[H(\bar{X}_{t_i}^n - y) \pi_{\bar{X}_{t_i}^n} [1](\bar{X}_{t_{i+1}}^n) \right]}.$$

où $\pi_{\bar{X}_{t_i}^n} [g]$ et $\pi_{\bar{X}_{t_i}^n} [1]$ sont donnés par (5.9).

2. Pour tous $k = 0, \dots, n-1$, et $y \in \mathbb{R}$, on a

$$\mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) | \bar{X}_{t_i}^n = y \right] = \frac{\mathbb{E} \left[H(\bar{X}_{t_i}^n - y) \pi_{\bar{X}_{t_i}^n}^\psi [g](\bar{X}_{t_{i+1}}^n) \right]}{\mathbb{E} \left[H(\bar{X}_{t_i}^n - y) \pi_{\bar{X}_{t_i}^n}^{\tilde{\psi}} [1](\bar{X}_{t_{i+1}}^n) \right]}.$$

où $\pi_{\bar{X}_{t_i}^n}^\psi [g]$ et $\pi_{\bar{X}_{t_i}^n}^{\tilde{\psi}} [1]$ sont donnés par (5.10).

Remarque 5.5. Les fonctions ψ et $\tilde{\psi}$ agissent comme des fonctions de localisation : on sélectionne les trajectoires qui ne sont pas trop éloignées de x . Typiquement, elles atteignent leur maximum en 0 et décroissent ensuite rapidement. On peut choisir les fonctions ψ et $\tilde{\psi}$ de manière à minimiser la variance intégrée des estimateurs du numérateur et dénominateur. Il est montré dans [12] que les fonctions ψ et $\tilde{\psi}$ optimales sont de type exponentielle $\psi(y) = e^{-\eta y}$, $\tilde{\psi}(y) = e^{-\tilde{\eta} y}$, où

$$\begin{aligned}\eta^2 &= \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n)^2 \pi_{\bar{X}_{t_i}^n}^2 [1](\bar{X}_{t_{i+1}}^n) \right] / \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n)^2 \right] \\ \tilde{\eta}^2 &= \mathbb{E} \left[\pi_{\bar{X}_{t_i}^n}^2 [1](\bar{X}_{t_{i+1}}^n) \right].\end{aligned}$$

On peut vérifier que η et $\tilde{\eta}$ ont une évolution avec n de l'ordre de \sqrt{n} .

Exercice 4. Démontrer un résultat analogue dans le modèle de Black et Scholes.

La formulation du Théorème 5.4 fournit un estimateur naturel pour l'espérance conditionnelle. On peut ainsi considérer N copies indépendantes $(\bar{X}^{n(j)})_{j=1}^N$ de \bar{X}^n et définir un estimateur de l'espérance conditionnelle par

$$\tilde{E} \left[g(\bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n \right] = \frac{\sum_{j=1}^N H(-\bar{X}_{t_i}^n + X_{t_i}^{n(j)}) \pi_{\bar{X}_{t_i}^n}^{\psi} [g](\bar{X}_{t_{i+1}}^{n(j)})}{H(-\bar{X}_{t_i}^n + X_{t_i}^{n(j)}) \pi_{\bar{X}_{t_i}^n}^{\psi} [1](\bar{X}_{t_{i+1}}^{n(j)})}$$

où $\delta_i^{(j)}$ et $\tilde{\delta}_i^{(j)}$ sont les opérateurs correspondant à la j -ème trajectoire simulée. En pratique ces estimateurs risquent d'être instables à cause de la division (ils n'ont d'ailleurs aucune raison d'être dans L^1 !). Il faut donc les corriger légèrement. Si ψ est à croissance polynômiale, il est assez facile de construire des polynômes $\underline{\rho}$ et $\bar{\rho}$ tels que

$$\underline{\rho}(\bar{X}_{t_i}^n) \leq \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n \right] \leq \bar{\rho}(\bar{X}_{t_i}^n).$$

On définit alors l'estimateur de l'espérance conditionnelle par

$$\hat{E} \left[g(\bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n \right] := \underline{\rho}(\bar{X}_{t_i}^n) \vee \tilde{E} \left[g(\bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n \right] \wedge \bar{\rho}(\bar{X}_{t_i}^n).$$

Comme $\bar{X}_{t_i}^n$ est dans tous les L^p , $p \geq 1$, on obtient ainsi un estimateur avec de bonnes propriétés d'intégrabilité. On a le résultat de convergence suivant.

Théorème 5.5. *Sous les hypothèses du Théorème 5.4, si on prend ψ et $\tilde{\psi}$ de la forme $e^{-\eta_n |x|}$ avec $\eta_n \sim \sqrt{n}$ quand $n \rightarrow \infty$, alors, pour tout $p \geq 1$,*

$$\left\| \hat{E} \left[g(\bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n \right] - \mathbb{E} \left[g(\bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n \right] \right\|_{L^p} \leq C \frac{n^{\frac{1}{4p}}}{N^{\frac{1}{2p}}}$$

où C ne dépend pas de n et N .

La preuve de ce résultat n'est pas très difficile mais un peu longue. Nous renvoyons à [13] pour les détails et pour l'étude du cas général.

Remarque 5.6. D'après le Lemme 5.1, si on connaît explicitement la densité $f_{\bar{X}_{t_i}^n}$ de $\bar{X}_{t_i}^n$, on peut également utiliser comme estimateur

$$\check{E} \left[g(\bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n = x \right] = \frac{\sum_{j=1}^N H(-\bar{X}_{t_i}^n + X_{t_i}^{n(j)}) \pi_{\bar{X}_{t_i}^n}^{\psi} [g](\bar{X}_{t_{i+1}}^{n(j)})}{f_{\bar{X}_{t_i}^n}(x)}.$$

Remarque 5.7. Lorsque $d \geq 2$, un résultat similaire à celui des Théorèmes 5.4 et 5.5 peut être obtenu mais l'écriture de l'estimateur est plus compliquée. Dans [12], vous trouverez une écriture en terme d'intégrales de Skorohod itérées. Le cas du modèle de Black-Scholes a été étudié par [20], la formulation est alors plus simple. On peut néanmoins obtenir une forme élégante de ce résultat dans le cas où $\bar{X}^n = X = W$. A un changement de variable près, il est souvent possible de se ramener à ce cas. On a alors sous des conditions analogues à celles du Théorème 5.4

$$\begin{aligned} & \mathbb{E} \left[g(W_{t_{i+1}}) \mid W_{t_i} = x \right] \\ &= \\ & \frac{\mathbb{E} \left[g(W_{t_{i+1}}) \prod_{\ell=1}^d \mathbf{1}_{x^\ell \leq W_{t_i}^\ell} e^{-\eta^\ell (W_{t_i}^\ell - x)} \left\{ \frac{n}{T} (\Delta^n W_i^\ell - \Delta^n W_{i+1}^\ell) + \eta^\ell \right\} \right]}{\mathbb{E} \left[\prod_{\ell=1}^d \mathbf{1}_{x^\ell \leq W_{t_i}^\ell} e^{-\eta^\ell (W_{t_i}^\ell - x)} \left\{ \frac{n}{T} \Delta^n W_i^\ell + \eta^\ell \right\} \right]} . \end{aligned}$$

La convergence de l'estimateur L^p est alors en $n^{\frac{d}{4p}}/N^{\frac{1}{2p}}$. Dans ce cas particulier, on peut simplifier la formulation de l'estimateur puisque l'on connaît explicitement la densité des W_{t_i} , voir Remarque 5.6.

Exemple 5.1. On considère le processus suivant

$$dX_t = \text{diag}[X_t] \sigma dW_t, \quad X_0^1 = X_0^2 = X_0^3 = 1.$$

avec

$$\sigma = \begin{bmatrix} 0.2 & 0 & 0 \\ 0.08 & 0.4 & 0 \\ 0.03 & -0.15 & 0.32 \end{bmatrix}.$$

On estime la densité du processus et l'espérance conditionnelle

$$r(x) = 100 \times \mathbb{E} \left[\left(\frac{X_2^1 + X_2^2}{2} - X_2^3 \right)^+ \mid X_1 = x \right]$$

en différents points. On donne le résultat moyen obtenu ainsi que l'écart-type de l'estimateur avec ou sans fonction de localisation.

		Densité $x^1 = 1.3$		
$x^3 \setminus x^2$		0.7	1.0	1.3
0.7	Valeur exacte	0.29	0.56	0.48
	Avec localisation	0.30[0.03]	0.56[0.02]	0.48[0.01]
	Sans localisation	0.30[0.41]	0.57[0.43]	0.51[0.44]
1.0	Valeur exacte	0.59	0.70	0.43
	Avec localisation	0.59[0.02]	0.70[0.01]	0.45[0.27]
	Sans localisation	0.60[0.47]	0.72[0.48]	0.47[0.49]
1.3	Valeur exacte	0.44	0.38	0.18
	Avec localisation	0.44[0.01]	0.38[0.00]	0.18[0.00]
	Sans localisation	0.45[0.48]	0.40[0.49]	0.22[0.51]

Espérance conditionnelle $x^1 = 0.9$

$x^3 \setminus x^2$		0.9	1.0	1.1
0.9	Valeur exacte	17.26	20.56	24.05
	Avec localisation	17.28[1.12]	20.49[1.19]	24.06[1.17]
	Sans localisation	16.88[5.01]	20.62[7.14]	25.04[11.52]
1.0	Valeur exacte	13.70	16.59	19.61
	Avec localisation	13.72[0.82]	16.59[0.92]	19.73[1.00]
	Sans localisation	12.97[6.20]	16.08[10.41]	21.35[25.48]
1.1	Valeur exacte	10.88	13.39	16.11
	Avec localisation	10.94[0.85]	13.48[0.90]	16.32[1.05]
	Sans localisation	10.58[17.54]	13.81[28.01]	13.19[166.22]

On observe une très forte volatilité de l'estimateur en l'absence de fonction de localisation. Il est donc absolument nécessaire de l'utiliser. Même avec celle-ci, la variance reste forte. Cette approche doit donc être combinée avec d'autres techniques de réduction de variance dès que cela est possible.

5.4 Application à l'évaluation d'options américaines

On revient maintenant au problème d'évaluation d'options américaines. On considère le problème discrétisé

$$\begin{aligned} \bar{p}^n(t_n, \bar{X}_{t_n}^n) &= g(\bar{X}_{t_n}^n) \\ \bar{p}^n(t_i, \bar{X}_{t_i}^n) &= \max \left\{ g(\bar{X}_{t_i}^n), e^{-rT/n} \mathbb{E} \left[\bar{p}^n(t_{i+1}, \bar{X}_{t_{i+1}}^n) \mid \bar{X}_{t_i}^n \right] \right\}, \quad i \in \{0, \dots, n-1\}. \end{aligned}$$

Si g est lipschitzienne, il est facile de construire une suite de polynômes $\bar{\rho}_i(x) = \alpha_i + \beta_i \|x\|^2$ tels que les (α_i, β_i) sont uniformément bornés en i et n et tels que

$$\begin{aligned} \left\| \mathbb{E} \left[g(\bar{X}_{t_n}^n) \mid \bar{X}_{t_{n-1}}^n \right] \right\| &\leq \bar{\rho}_{n-1}(\bar{X}_{t_{n-1}}^n) \\ \text{et } \forall i \quad \left\| \mathbb{E} \left[g(\bar{X}_{t_i}^n) \vee e^{-rT/n} \bar{\rho}_i(\bar{X}_{t_i}^n) \mid \bar{X}_{t_{i-1}}^n \right] \right\| &\leq \bar{\rho}_{i-1}(\bar{X}_{t_{i-1}}^n). \end{aligned}$$

On peut donc utiliser l'estimateur du Théorème 5.5 pour calculer les espérances conditionnelles. L'algorithme utilisé est le suivant :

On considère nN copies $(\bar{X}^{n(1)}, \dots, \bar{X}^{n(nN)})$ de $\bar{X}^n =: \bar{X}^{n(0)}$, et on pose $\mathcal{N}_i := \{(i-1)N + 1, \dots, iN\}$.

Initialisation : Pour $j \in \{0\} \cup \mathcal{N}_n$, on pose : $\hat{p}^n(t_n, \bar{X}_{t_n}^{n(j)}) = g(\bar{X}_{t_n}^{n(j)})$.

Récurrence rétrograde : Pour $i = n, \dots, 1$, on pose, pour $j \in \{0\} \cup \mathcal{N}_{i-1}$:

$$\hat{p}^n(t_{i-1}, \bar{X}_{t_{i-1}}^{n(j)}) = \max \left\{ g(\bar{X}_{t_{i-1}}^{n(j)}), e^{-rT/n} \hat{E} \left[\hat{p}^n(t_i, \bar{X}_{t_i}^{n(j)}) \mid \bar{X}_{t_{i-1}}^{n(j)} \right] \right\}$$

où

$$\begin{aligned} \hat{E} \left[\hat{p}^n(t_i, \bar{X}_{t_i}^{n(j)}) \mid \bar{X}_{t_{i-1}}^{n(j)} \right] &= -\bar{\rho}_{i-1}(\bar{X}_{t_{i-1}}^{n(j)}) \vee \tilde{E} \left[\hat{p}^n(t_i, \bar{X}_{t_i}^{n(j)}) \mid \bar{X}_{t_{i-1}}^{n(j)} \right] \wedge \bar{\rho}_{i-1}(\bar{X}_{t_{i-1}}^{n(j)}) \\ \tilde{E} \left[\hat{p}^n(t_i, \bar{X}_{t_i}^{n(j)}) \mid \bar{X}_{t_{i-1}}^{n(j)} \right] &:= \frac{\sum_{\ell \in \mathcal{N}_i} \mathbf{1}_{\bar{X}_{t_{i-1}}^{n(j)} \leq \bar{X}_{t_{i-1}}^{n(\ell)}} \hat{p}^n(t_i, \bar{X}_{t_i}^{n(\ell)}) \delta_{i-1}^{(\ell)} \left(\psi(\bar{X}_{t_{i-1}}^{n(\ell)} - \bar{X}_{t_{i-1}}^{n(j)}) \right)}{\sum_{\ell \in \mathcal{N}_i} \mathbf{1}_{\bar{X}_{t_{i-1}}^{n(j)} \leq \bar{X}_{t_{i-1}}^{n(\ell)}} \tilde{\delta}_{i-1}^{(\ell)} \left(\tilde{\psi}(\bar{X}_{t_{i-1}}^{n(\ell)} - \bar{X}_{t_{i-1}}^{n(j)}) \right)} \end{aligned}$$

pour $i \geq 1$ et pour $i = 1$

$$\hat{E} \left[\hat{p}^n(t_1, \bar{X}_{t_1}^{n(j)}) \right] := \frac{1}{N} \sum_{\ell \in \mathcal{N}_1} \hat{p}^n(t_1, \bar{X}_{t_1}^{n(\ell)}) .$$

Comme on fait n approximations de l'espérance conditionnelle, l'erreur est de l'ordre de n fois celle donnée par le Théorème 5.5, i.e.

$$\max_{0 \leq i \leq n} \|\hat{p}^n(t_i, \bar{X}_{t_i}^n) - \bar{p}^n(t_i, \bar{X}_{t_i}^n)\|_{L^p} \leq Cn \left(n^{1/4p} / N^{1/2p} \right) .$$

Remarque 5.8. *On pourra bien évidemment, et c'est même fortement recommandé, coupler cet algorithme avec une méthode de réduction de variance de type variable de contrôle.*

Remarque 5.9. Tout ceci se généralise à une large classe d'équations forward-backward. La présentation générale de la méthode et les vitesses de convergence sont dues à [13]. Lorsque le processus \bar{X}^n est de dimension d , la vitesse de convergence devient $n \left(n^{d/4p} / N^{1/2p} \right)$. Cette approche a été initiée par [19] et [43] mais en utilisant des outils différents pour le calcul des espérances conditionnelles.

Bibliographie

- [1] P. Baldi, L. Caramellino, and M.G. Iovino. Pricing single and double barrier options via sharp large deviations. *Preprint*, 1997.
- [2] P. Baldi, L. Caramellino, and M.G. Iovino. Pricing complex barrier options with general features using sharp large deviation estimate. *Proceedings of the MCQMC Conference, Calremont (LA), USA*, 1999.
- [3] P. Baldi, L. Caramellino, and M.G. Iovino. Pricing general barrier options : a numerical approach using sharp large deviations. *To appear in Mathematical Finance (1999)*, 1999.
- [4] V. Bally, L. Caramellino, and A. Zanette. Pricing and hedging american options by monte carlo using a malliavin calculus approach. Technical Report 4804, INRIA, 2003.
- [5] V. Bally and G. Pagès. A quantization algorithm for solving multi-dimensional optimal stopping problems. *Technical report 628, université Paris 6*, 2000.
- [6] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations (I) : convergence rate of the distribution function. *Probability Theory and Related Fields*, 104(1), 1996.
- [7] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations (II) : convergence rate of the density. *Monte Carlo Methods and Applications*, 2 :93–128, 1996.
- [8] J. Barraquand and D. Martineau. Numerical valuation of high dimensional multivariate american securities. *J. of Finance and Quantitative Analysis*, 30 :383–405, 1995.
- [9] E. Benhamou. An application of malliavin calculus to continuous time asian options greeks. Technical report, Financial Markets Research Centre, London Business School, 2000.
- [10] E. Benhamou. Faster greeks for discontinuous payoff options (a malliavin calculus approach in black world. Technical report, Financial Markets Research Centre, London Business School, 2000.
- [11] E. Benhamou. A generalisation of malliavin weighted scheme for fast computation of the greeks. Technical report, Financial Markets Research Centre, London Business School, 2000.
- [12] B. Bouchard, I. Ekeland, and N. Touzi. On the malliavin approach to monte carlo approximation of conditional expectations. Technical report, preprint, 2002.
- [13] B. Bouchard and N. Touzi. Discrete time approximation and monte-carlo simulation of backward stochastic differential equations. Technical report, preprint, 2002.

- [14] M. Broadie and P. Glassermann. Pricing american-style securities using simulation. *J. of Economic Dynamics and Control*, 21 :1323–1352, 1997.
- [15] M. Broadie and P. Glassermann. A stochastic mesh method for pricing high-dimensional american options. *Working Paper*, Columbia University :1–37, 1997.
- [16] J.M.C. Clark and R.J. Cameron. The maximum rate of convergence of discrete approximations for stochastic differential equations. In B. Grigelionis, editor, *Stochastic Differential Systems – Filtering and Control*, volume 25 of *Lecture Notes in Control and Information Sciences*. Proceedings of the IFIP Working Conference, Vilnius, Lithuanie, 1978, Springer-Verlag, 1980.
- [17] E. Clement, D. Lamberton, and P. Protter. An analysis of the lonstaff-schwartz algorithm for american option pricing. Technical Report 04/2001, University of Marne-La-Vallée, April 2001.
- [18] J. Cvitanič, J. Ma, and J. J. Zhang. Efficient computation of hedging portfolios for options with discontinuous payoffs. *to appear in Mathematical Finance*, 2001.
- [19] Carriere E. Valuation of the early-exercise price for options using simulations and non-parametric regression. *Insurance : mathematics and Economics*, 19 :19–30, 1996.
- [20] Lions P.-L. et H. Regnier. Calcul du prix et des sensibilités d’une option américaine par une méthode de monte carlo. Technical report, preprint, 2001.
- [21] Kohatsu-Higa A. et R. Pettersson. Variance reduction methods for simulation of densities on wiener space. *SIAM Journal on Numerical Analysis*, 2001.
- [22] O. Faure. Numerical pathwise approximation of stochastic differential equation. *Applied Stochastic Models and Data Analysis*, 1992.
- [23] O. Faure. *Simulation du Mouvement Brownien et des Diffusions*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 1992.
- [24] E. Fournié, J. M. Lasry, J. Lebuchoux, and P. L. Lions. Applications of malliavin calculus to monte-carlo methods in finance. i. *Finance and Stochastics*, 3(4) :391–412, 1999.
- [25] E. Fournié, J. M. Lasry, J. Lebuchoux, and P. L. Lions. Applications of malliavin calculus to monte-carlo methods in finance. ii. *Finance and Stochastics*, 5(2) :201–236, 2001.
- [26] J.G. Gaines and T.J. Lyons. Variable step size control in the numerical solution of stochastic differential equations. (submitted), 1991.
- [27] Paul Glasserman and David D. Yao. Some guidelines and guarantees for common random numbers. *Manage. Sci.*, 38(6) :884–908, 1992.
- [28] P.W. Glynn. Optimization of stochastic systems via simulation. In *Proceedings of the 1989 Winter simulation Conference*, pages 90–105. San Diego, Society for Computer Simulation, 1989.
- [29] A. Gobet, E. et Kohatsu-Higa. Computation of greeks for barrier and lookback options using malliavin calculus. *Mathematical Finance*, 2001.
- [30] E. Gobet. Schema d’Euler continu pour des diffusions tuees et options barriere. (Continuous Euler scheme for killed diffusions and barrier options). *C. R. Acad. Sci., Paris, Ser. I, Math.*, 326(12) :1411–1414, 1998.
- [31] E. Gobet. Schema d’Euler discret pour diffusion multidimensionnelle tuee. (Discrete Euler scheme for killed multidimensional diffusion). *C. R. Acad. Sci., Paris, Ser. I, Math.*, 328(6) :515–520, 1999.

- [32] E. Gobet. Weak approximation of killed diffusion using euler schemes. *Stochastic Processes and their Applications*, 87 :167–197, 2000.
- [33] E. Gobet. Euler schemes and half-space approximation for the simulation of diffusion in a domain. *ESAIM Probability and Statistics*, 5 :261–297, 2001.
- [34] E. Gobet and R. Munos. Sensitivity analysis using itô-malliavin calculus and martingales. Technical report, preprint CMAP, école Polytechnique, 2002.
- [35] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes (2nd Ed.)*. North Holland, 1989.
- [36] S. Kanagawa. On the rate of convergence for Maruyama’s approximate solutions of stochastic differential equations. *Yokohama Math. J.*, 36 :79–85, 1988.
- [37] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, second edition, 1991.
- [38] A.G.Z Kemna and A.C.F. Vorst. A pricing method for options based on average asset values. *J. Banking Finan.*, pages 113–129, March 1990.
- [39] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer Verlag, 1992.
- [40] B. Lapeyre, E. Pardoux, and R. Sentis. *Méthodes de Monte Carlo pour les équations de transport et de diffusion*. Number 29 in Mathématiques & Applications. Springer-Verlag, 1998.
- [41] B. Lapeyre and E. Temam. Competitive monte carlo methods for asian options. *Journal of Computational Finance*, 5, 2001.
- [42] P. L’Ecuyer and G. Perron. On the convergence rates of ipa and fdc derivatives estimators. *Oper. Res.*, 42 :643–656, 1994.
- [43] F.A. Longstaff and E.S. Schwartz. Valuing american options by simulations :a simple least-squares approach. *Working Paper Anderson Graduate School of Management University of California*, 25, 1998.
- [44] Annie Miller. Méthodes de monte carlo. Note de Cours.
- [45] N.J. Newton. Variance reduction for simulated diffusions. *SIAM J. Appl. Math.*, 54(6) :1780–1805, 1994.
- [46] D. Nualart. *The Malliavin Calculus and Related Topics*. Springer Verlag, Berlin, 1995.
- [47] D. Nualart. Analysis on wiener space and anticipating stochastic calculus. *Lect. Notes Math. 1690, Springer, Berlin*, 1998.
- [48] B. Øksendal. An introduction to malliavin calculus with applications to economics. Technical report, Lecture Notes from a course given 1996 at the Norwegian School of Economics and Business Administration (NHH), NHH Preprint Series, 1996.
- [49] G. Pages, V. Bally, and Printems. A stochastic optimisation method for nonlinear problems. *Technical report number 02, University Paris XII*, 2001.
- [50] W. H. Press, Saul A. Teukolsky, W. T. Vetterling, and Brian P. Flannery. *Numerical recipes in C*. Cambridge University Press, Cambridge, second edition, 1992. The art of scientific computing.

- [51] P. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin, 1990.
- [52] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin, 1991.
- [53] L. C. G. (ed.) Rogers and al., editors. *Monte Carlo methods for stochastic volatility models*. Cambridge University Press., 1997.
- [54] P. Seumen Tonou. *Méthodes Numériques Probabilistes pour la résolution d'équations de transport et pour l'évaluation d'options exotiques*. PhD thesis, Université de Provence, 1997.
- [55] D. Talay. Discrétisation d'une e.d.s. et calcul approché d'espérances de fonctionnelles de la solution. *Mathematical Modelling and Numerical Analysis*, 20(1) :141–179, 1986.
- [56] D. Talay. Simulation and numerical analysis of stochastic differential systems : a review. In P. Kree and W. Wedig, editors, *Probabilistic Methods in Applied Physics*, volume 451 of *Lecture Notes in Physics*, chapter 3, pages 54–96. Springer-Verlag, 1995.
- [57] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4) :94–120, 1990.
- [58] E. Temam. *Couverture d'options exotiques-Pricing des options asiatiques*. PhD thesis, Université Paris VI, 2001.
- [59] J.N. Tsitsiklis and B. Van Roy. Regression methods for pricing complex american-style options. *Working Paper*, MIT :1–22, 2000.